Tensor product methods and entanglement optimization for ab initio quantum chemistry

Szilárd Szalay* Max Pfeffer† Valentin Murg‡ Gergely Barcza*
Frank Verstraete‡ Reinhold Schneider† Örs Legeza*

December 19, 2014

Abstract

The treatment of high-dimensional problems such as the Schrödinger equation can be approached by concepts of tensor product approximation. We present general techniques that can be used for the treatment of high-dimensional optimization tasks and time-dependent equations, and connect them to concepts already used in many-body quantum physics. Based on achievements from the past decade, entanglement-based methods, – developed from different perspectives for different purposes in distinct communities already matured to provide a variety of tools – can be combined to attack highly challenging problems in quantum chemistry. The aim of the present paper is to give a pedagogical introduction to the theoretical background of this novel field and demonstrate the underlying benefits through numerical applications on a text book example. Among the various optimization tasks we will discuss only those which are connected to a controlled manipulation of the entanglement which is in fact the key ingredient of the methods considered in the paper. The selected topics will be covered according to a series of lectures given on the topic *"New wavefunction methods and entanglement optimizations in quantum chemistry"* at the Workshop on Theoretical Chemistry, 18 - 21 February 2014, Mariapfarr, Austria.

Contents

*Strongly correlated systems "Lendület" research group, Wigner Research Centre for Physics, Konkoly-Thege Miklós út 29-33, 1121 Budapest, Hungary

†Fakultät II - Mathematik und Naturwissenschaften, Institut für Mathematik, Technische Universität Berlin, Strasse des 17. Juni 136, Berlin, Germany

‡Fakultät für Physik, Universität Wien, Boltzmanngasse 3, A-1090 Vienna, Austria

1 Introduction

For the approximation of the wave function of the electronic structure of an atomic or molecular system, any method chosen will have to compromise between the demanded accuracy on the one hand and the high computational complexity of the task on the other. While *Density Functional Theory* (DFT)[47] and *Coupled Cluster* (CC) or *Quantum Monte Carlo* methods[96,204,226] are in this sense standard methods for the quantitative study of large weakly correlated systems, there has been no method-of-choice solution for finding a sufficiently accurate, data-sparse representation of the exact many-body wave function if many electrons are *strongly correlated,* as, for instance, in open-shell systems as transition metal complexes[15,26,30,53,94,114,148,173,194,198,199,217,261].

Due to the many-electron interactions present, strongly correlated problems cannot be sufficiently described by small perturbations of a single Slater determinant. For the treatment of other many-particle systems, e.g., spin systems, alternative representations have been proposed, resulting in the development of so-called *Matrix Product States* (MPS)[184,211,212,238]. The MPS method represents the wavefunction of a system of d components or *"sites"* (corresponding, e.g., to *molecular orbitals*) by forming products of d matrices, each belonging to one component of the system. The computational complexity of the task is now governed by the size of these matrices, related to the eigenvalue spectrum of the corresponding subsystem density matrices[187] characterizing in a formal way the so-called *entanglement* among the different components[11,64,95,125,127,193,234,241]. MPS consists in a linear arrangement of the components, while more recently the approach has been generalized to so-called *Tensor Network States* (TNS)[45,67,69,135,136,147,151,167,168,180,219,236,237,240], allowing a more flexible connection of the components of the respective system. Identical, but independent approaches were devised in numerical mathematics under the term of tensor product approximation, where low-rank

factorization of matrices is generalized to higher order tensors[82–84]. In quantum chemistry, the MPS[35,36,39–41,43,57,58,74,79,80,114,125–127,129,144,148,150,158,160,162–165,176,201,253,268,269,274–276 15,26,27,29,30,62,70,94,108,109,112,113,115,117,139,146,147,151,157,161,166,173,206,217,218,229,261,262,264,270] and TNS[147,169,170,172] representation can be used to approximate the full-CI wave function[37,38,42,44,111,124,135,136,149,151,259,265]. By this new concept of data-sparse representation, an accurate representation of the electronic structure will then be possible in polynomial time if the exact wave function can be approximated to a sufficient extent by moderately entangled TNS representations. The underlying *Molecular Orbital* (MO) basis can be optimized by well known techniques from multi-configurational methods[96] as, e.g., *Multi Configuration Self Consistent Field* (MCSCF) method, which constitutes a tensor approximation method as well at the level of first quantization.

Entanglement-based methods, – developed from different perspectives for different purposes in distinct communities, already matured to provide a variety of tools – can be combined to attack highly challenging problems in quantum chemistry[1–5]. A very promising direction is, especially, to develop and implement an efficient *Quantum Chemistry algorithm based on Tree Tensor Network States* (QC-TTNS), in particular enabling the treatment of problems in quantum chemistry that are intractable by standard techniques as DFT or CC[135,169,170,172].

The aim of the present paper is to give a pedagogical introduction to the theoretical background of this novel field and demonstrate the underlying benefits through numerical applications on a text book example. We give a technical introduction to low rank tensor factorization and do not intend to present a detailed review of the field. Only some selected topics will be covered according to lectures given on the topic *"New wave function methods and entanglement optimizations in quantum chemistry"* at the Workshop on Theoretical Chemistry, 18 - 21 February 2014, Mariapfarr, Austria[6]. In accordance with this, the organization of the present paper is as follows. In Secs. 2 and 3 a very detailed description of the theory follows so that those interested readers who just entered in the field could follow recent developments. A brief summary in order to highlight the most important concepts used in numerics is presented in Sec. 4 together with numerical applications by outlining ideas and existing algorithmic structures that have been used to arrive at an efficient implementation. At this stage among the various optimization tasks only those will be analyzed which are connected directly to the manipulation of entanglement, which is in fact the key ingredient of the methods presented in the paper. To unify notations, in what follows, we mainly use terms and notations as common in physics and chemistry.

1.1 Tensor product methods in quantum chemistry

Multi-particle Schrödinger-type equations constitute an important example of problems posed on high-dimensional tensor spaces. Numerical approximation of solutions of these problems suffers from the *curse of dimensionality*, i.e., the computational complexity scales exponentially with the dimension of the space. Circumventing this problem is a challenging topic in modern numerical analysis with a variety of applications, covering aside from the electronic and nuclear Schrödinger equation e.g., the Fokker-Planck equation and the chemical master equation[141]. Considerable progress in the treatment of such problems has been made by concepts of tensor product approximation[81–83].

In the year 1992, S. R. White introduced a very powerful numerical method, the Density-Matrix Renormalisation Group (DMRG)[248–250]. It allows us to determine the physical properties of low-dimensional correlated systems such as quantum spin chains or chains of interacting itinerant electrons to unprecedented accuracy[61,92,93,178,186,211].

Further success of the DMRG method in quantum physics motivated its application to *Quantum Chemical* problems (QC-DMRG)[39,125,159,162,253]. In the treatment of problems where large active spaces are mandatory to obtain reliable and accurate results, it has proven capable of going well beyond the limits of present day quantum chemistry methods and even reach the full-CI limit[37,38,42,44,111,124,135,136,149,151,259,260].

In the past decade, the method has gone through major algorithmic developments by various groups[38,111,124,150,150,211,265]. For example, the two-body reduced density matrices calculated with the DMRG method[39,125,275] can be used in the standard orbital optimization procedure[96]. Resulting methods are the *DMRG Complete Active Space Self Consistent Field* (DMRG-CASSCF) or *DMRG Self Consistent Field* (DMRG-SCF)[74,268,274,275]. Another direction is the post-DMRG treatment of dynamic correlation. DMRG as it can be considered as a *CAS Configuration Interaction* (CAS-CI) technique can recover static correlation, and, depending on the size of the active space, one can afford also some portion of the dynamic correlation. Quite recently, various advanced methods accounting for dynamic correlation on top of the DMRG framework have been developed[115,139,176,206,251,269,270]. The first implementation of the relativistic quantum chemical *two- and four-component* density matrix renormalization group algorithm (2c- and 4c-DMRG) has also been presented[109].

The DMRG method can be used to calculate ground as well as excited states. This can be achieved either by targeting several of them in a state average fashion[57,74,125,126,139,165,170,218,264] or alternatively based on the MPS tangent vectors[88,89,173,190,263]. Since the DMRG method is very flexible it can be used even in such situations when the wave function character changes dramatically[29,126,165,170,264]. Additionally, the ansatz is size consistent by construction and symmetries as particle number, spin projection[253], spin reflection symmetries[137], Abelian point group symmetries[41,126,127] and even non-Abelian symmetries can be factored out explicitly[123,152–155,192,205,217,220–223,230,231,245,245,262,264,276]. Quite recently, MPS and further tensor product approximations have been applied in *post Hartree-Fock* (post-HF) methods to the decomposition of the two electron integrals, the AO-MO *(Atomic Orbital-Molecular Orbital)* transformation, and the *Møller-Plesset perturbation theory* (MP2) energy expression[22].

In the MPS like methods the computational complexity of the task is governed by the size of the matrices used to approximate the wavefunction, which can, however, be controlled based on various truncation criteria to achieve a priory set error margin[99,125]. In a system with identical sites, this feature is directly connected to the scaling of entanglement when subsystems include larger and larger portion of the total system, also called as area law[64,95,193,234]. The situation is more complicated in quantum chemical applications since the ranks of the matrices also depend strongly on the ordering of the matrices[15,30,125–127], thus different orderings lead to better or worse results if the ranks are kept fixed[39,114,146,158,161,162,201,264,269]. Another main aspect that effects the performance of the method is the optimization of the basis[74,133,146,169,268,274] and initialization of the network[15,39,127,159,163,264]. Even though the significant efforts dedicated to the various optimization tasks, it remains an open question to determine the minimum of computational effort to obtain results with a given accuracy threshold.

Shortly after DMRG was introduced, it was found that DMRG may also be phrased in terms of MPS[184], first formulated for special spin systems as the *Affleck-Kennedy-Lieb-Tasaki* (AKLT) model[8]. More recently, the *Higher Order Singular Value Decomposition* (HOSVD)[237,238,243] have made MPS the basis of variational frameworks and revealed a profound connection to quantum information theory[127,212,241]. In this context, it became apparent that MPS is only one of a more general set of formats: while MPS corresponds to an arrangement of orbitals in a linear topology, quantum states may more generally be arranged as more complex topologies, leading to tensor network states TNS[147,169,170,172]. For applications to smaller systems, prototypical tensor-network state approaches to quantum chemistry have already been developed, including the so called *Complete Graph Tensor Network State* (CGTNS) approach[147], and the *Tree Tensor Network State* (TTNS) approach[169,170,172]. The QC-TTNS combines a number of favorable features that suggest it might represent a novel, flexible approach in quantum chemistry: the more general concept of data-sparsity inherent in the TNS representation allows for the efficient representation of a much bigger class of wave functions than accessible by state-of-the-art methods. The desired accuracy may be adjusted, so that the ansatz in principle permeates the whole full-CI space.

These developments foster the hope that with their help some of the major questions in quantum chemistry and condensed matter physics may be solved. The concept of MPS and tree structured tensor network states has been rediscovered independently in numerical mathematics for tensor product approximation[81,183].

1.2 Entanglement and quantum information entropy in quantum chemistry

In quantum systems, correlations having no counterpart in classical physics arise. Pure states showing these strange kinds of correlations are called entangled ones[11,64,95,103,193,195,227,234], and the existence of these states has so deep and important consequences[21,48,63] that Schrödinger has identified entanglement to be the characteristic trait of quantum mechanics[214,215]. The QC-DMRG and QC-TTNS algorithms approximate a composite system with strong interactions between many pairs of orbitals, and it turned out that the results of quantum information theory[177,257] can be used to understand the criteria of their convergence.[127,211,212,238,241].

Recently, quantum information theory has also appeared in quantum chemistry giving a fresh impetus to the development of methods in electronic structure theory[127,129,171,174,201,277 9,15,17,26,28–30,62,70,105,109,112,145,166]. The amount of contribution of an orbital to the total correlation can be characterized, for example, by the *single-orbital entropy*[127], and the the sum of all single-orbital entropies gives the amount of *total correlation* encoded in the wave function[129,133]. This quantity can be used to monitor changes in entanglement as system parameters are adjusted, for example, changing bond length or other geometrical properties[62,70,170]. A useful quantity to numerically characterize the correlations (classical and quantum together) between pairs of orbitals is the *mutual information*[15,130,201] and it together with the orbital entropy provides chemical information about the system, especially about bond formation and nature of static and dynamic correlation[15,29,30,62].

The two-orbital mutual information also yields a weighted graph of the overall two-orbital correlation of both classical and quantum origin reflecting the entanglement topology of the

molecules. Therefore, this quantity can also be used to carry out optimization tasks based on the entanglement between the different components – itself determining the complexity of the computation – since it depends strongly on the chosen network topology and is in principle unknown for a given system. To promote the efficiency of tensor product methods various entanglement-based approaches are used to determine, for example, the appropriate ordering, network topology, optimal basis, and efficient network initialization. These important and non-trivial tasks will be considered in Sec. 4.

1.3 Tensor decomposition methods in mathematics

Approximation of quantities in high dimensional spaces is a hard problem with a variety of applications, and the development of generic methods that circumvent the enormous complexity of this task have recently, independent of the developments in the study of quantum systems, gained significant interest in numerical mathematics[82]. A recent analysis shows that, beyond the matrix case (corresponding to tensors of order 2), almost all tensor problems, even that of finding the best rank-1 approximation of a given tensor, are in general NP hard[97]. Although this shows that tensor product approximation in principle is an extremely difficult task, a variety of generic concepts for the approximation of solutions of certain problem classes have recently been proposed[83,84], some of which[56,78,82,140,141,203,232] bear a surprising similarity to methods used to treat problems in quantum physics[86,87].

The classical *Tucker format* attains sparsity via a subspace approximation. Multiconfigurational methods like MCSCF or CASSCF are in fact a Tucker approximation in the framework of antisymmetry. Its unfavorable scaling has recently been circumvented by a multilevel or hierarchical subspace approximation framework named *Hierarchical Tucker format*[81,82], interestingly corresponding to the TTNS. A similar format called *Tensor Trains,* developed independently[181,182,208], is a formal version of the MPS with open boundary conditions. Investigation of the theoretical properties of TNS and MPS in a generic context have shown that they inherit desirable properties of matrix factorization. E.g., closedness of the set of tensors of fixed block size[82] implies the existence of minimizers in these sets for convex optimization problems. Also, these sets possess a manifold structure that helps to remove redundancy in the parametrization by the introduction of so-called gauge conditions[100]. They can be used to set up variational frameworks for the treatment of optimization problems[135] and of time-dependent problems[56,87,140,143], bearing again a close connection to approaches in the quantum physics community[86]. In this general context, the robustness and quasi-best approximation of the HOSVD, studied in the mathematics community[76,82,83], and of the (one site) DMRG[252] as simple and very efficient numerical methods for the treatment of optimization problems are now well-understood[82,84]. These fundamental properties establish MPS and TTNS as advantageous concepts in tensor product approximation. It is important to note that all these properties are no longer valid in general if the tensor networks contains closed loops, as in case of the projected entangled pair states (PEPS)[236] and the multiscale entanglement renormalization ansatz (MERA)[239]. It is still widely unexplored under which conditions the favorable properties of tree structured TNS can be extended to general TNS. In mathematics the phrases *hierarchical tensor representation* or *Hierarchical Tucker format* as well as Tensor Trains instead of MPS are used since there the focus is not only on quantum mechanical systems, but rather on universal tools to handle high-dimensional

approximations. Many of the recent developments in mathematics parallel others in quantum computations on a more formal, generic level, often faced with similar experiences and similar problems.

2 Quantum chemistry

2.1 The electronic Schrödinger equation

A quantum mechanical system of N non-relativistic electrons is completely described by a *state-function* Ψ depending on $3N$ spatial variables $\mathbf{r}_a \in \mathbb{R}^3$, $a = 1, \ldots, N$, together with N discrete spin variables $s_a \in \{\pm\frac{1}{2}\}$, $a = 1, \ldots, N$,

$$\begin{aligned} \Psi : \mathbb{R}^{3N} \otimes \left\{\pm\frac{1}{2}\right\}^N \cong \left(\mathbb{R}^3 \otimes \left\{\pm\frac{1}{2}\right\}\right)^N \quad &\longrightarrow \quad \mathbb{C} \\ (\mathbf{r}_1, s_1; \ldots; \mathbf{r}_N, s_N) \quad &\longmapsto \quad \Psi(\mathbf{r}_1, s_1; \ldots; \mathbf{r}_N, s_N). \end{aligned} \tag{1}$$

The function Ψ belongs to the Hilbert space $L_2\big((\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N\big)$ having the standard inner product

$$\langle \Psi, \Phi \rangle = \sum_{s_i = \pm\frac{1}{2}} \int_{\mathbb{R}^{3N}} \overline{\Psi(\mathbf{r}_1, s_1; \ldots; \mathbf{r}_N, s_N)} \Phi(\mathbf{r}_1, s_1; \ldots; \mathbf{r}_N, s_N) \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N, \tag{2}$$

and the norm $\|\Psi\| = \sqrt{\langle \Psi, \Psi \rangle}$. The *Pauli antisymmetry principle* states that the wave function of *fermions,* in particular electrons, must be antisymmetric with respect to the permutation of variables, i.e., for $a \neq b$

$$\Psi(\ldots; \mathbf{r}_a, s_a; \ldots; \mathbf{r}_b, s_b; \ldots) = -\Psi(\ldots; \mathbf{r}_b, s_b; \ldots; \mathbf{r}_a, s_a; \ldots). \tag{3}$$

Such wave-functions are the elements of the antisymmetric tensor subspace $\bigwedge_{i=1}^N L_2\big(\mathbb{R}^3 \times \{\pm\frac{1}{2}\}\big)$. The *Pauli exclusion principle* immediately follows: Ψ must vanish for the points of $(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N$ which have the coordinates $\mathbf{r}_a = \mathbf{r}_b$ and $s_a = s_b$ for some $a \neq b$ fermions[47,197].

In quantum mechanics, we are usually interested in wave-functions having definite energies. This is expressed by the *stationary Schrödinger equation*,

$$H\Psi = E\Psi, \tag{4}$$

i.e., the wave function is an eigenfunction of a differential operator, namely the *Hamilton operator* H, and the eigenvalue $E \in \mathbb{R}$ is the energy of the corresponding state Ψ. One of the most important quantities is the *ground state energy* E_0, which is the lowest eigenvalue. The well known *Born-Oppenheimer-approximation* considers a nonrelativistic quantum mechanical system of N electrons in an exterior field generated by the K nuclei. In this case H is as follows

$$H = H_{\mathrm{kin}} + H_{\mathrm{pot}}, \qquad H_{\mathrm{pot}} = H_{\mathrm{ext}} + H_{\mathrm{int}}, \tag{5a}$$

$$H_{\mathrm{kin}} = \sum_{a=1}^{N} -\frac{1}{2}\Delta_a, \qquad H_{\mathrm{ext}} = -\sum_{a=1}^{N}\sum_{c=1}^{K} \frac{Z_c}{|\mathbf{R}_c - \mathbf{r}_a|}, \qquad H_{\mathrm{int}} = \frac{1}{2}\sum_{\substack{a,b=1 \\ b\neq a}}^{N} \frac{1}{|\mathbf{r}_b - \mathbf{r}_a|}. \tag{5b}$$

Since the Hamilton operator is a linear second order differential operator, the analysis for the electronic Schrödinger equation has been already established to a certain extent. We would like to briefly summarize some basic results and refer to the literature[47,197]. The *Sobolev spaces* $H^m := H^m\big((\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N\big)$, $m \in \mathbb{N}_0$ are defined as the spaces of functions for which all derivatives up to order m are in $H^0 := L_2\big((\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N\big)$. Consequently, the operator H maps the Sobolev space H^1 continuously into its dual space H^{-1}, i.e., $H : H^1 \to H^{-1}$ boundedly[47,197,271]. The potential operator H_{pot} maps the Sobolev spaces H^1 continuously into H^0, i.e., $H_{\text{pot}} : H^1 \to H^0 = L_2$ boundedly[197,271]. The electronic Schrödinger operator admits a rather complicated spectrum. We are interested mainly in the ground state energy E_0. If $\sum_{c=1}^K Z_c \geq N$, in particular for electrical neutral systems, it is known[197,272] that E_0 is an eigenvalue of finite multiplicity of the operator $H : H^2 \to H^0$ below the essential spectrum $\sigma_{\text{ess}}(H)$ of H, i.e., $-\infty < E_0 < \inf \sigma_{\text{ess}}(H)$. Summing up, the energy space for the electronic Schrödinger equation is

$$\mathcal{V}_N = H^1\Big(\Big(\mathbb{R}^3 \times \Big\{\pm\frac{1}{2}\Big\}\Big)^N\Big) \cap \bigwedge_{i=1}^N L_2\Big(\mathbb{R}^3 \times \Big\{\pm\frac{1}{2}\Big\}\Big). \tag{6}$$

This situation will be considered in the sequel.

For the sake of simplicity, we will also always assume that E_0 is a simple eigenvalue, i.e., of multiplicity one. In the case we deal with here, i.e., the stationary electronic Schrödinger equation in non-relativistic and Born-Openheimer setting, we can assume without the loss of generality that the wave function is real valued. (This does not hold for linear response theory or time-dependent problems, as well as for the relativistic regime, where complex phases play an important role.) According to the well known mini-max principle[197], the ground state energy and the corresponding wave function satisfies the *Rayleigh-Ritz variational principle*[197,271], i.e., the lowest eigenvalue is the minimum of the Rayleigh quotient $\frac{\langle \Psi, H\Psi\rangle}{\langle \Psi, \Psi\rangle}$, or equivalently,

$$E_0 = \min\big\{\langle \Psi, H\Psi\rangle : \langle \Psi, \Psi\rangle = 1, \Psi \in \mathcal{V}_N\big\}, \tag{7a}$$

$$\Psi_0 = \operatorname{argmin}\big\{\langle \Psi, H\Psi\rangle : \langle \Psi, \Psi\rangle = 1, \Psi \in \mathcal{V}_N\big\}. \tag{7b}$$

Since the Hamilton operator maps $H : \mathcal{V}_N \to (\mathcal{V}_N)^*$ boundedly, we will put the eigenvalue problem into the following weak formulation[271], to find the normalized $\Psi_0 \in \mathcal{V}_N$, satisfying

$$\langle \Phi, (H - E_0)\Psi_0\rangle = 0, \qquad \langle \Psi_0, \Psi_0\rangle = 1, \qquad \forall \Phi \in \mathcal{V}_N. \tag{8}$$

We will consider the above framework[47,272] throughout the present paper.

2.2 Full configuration interaction approach and the Ritz-Galerkin approximation

A convenient way to approximate the wave function is to use an anti-symmetric tensor product of basis functions depending only on single particle variables $(\mathbf{r}_a, s_a)$, which can be realized by determinants. To this end, let us consider a finite subset of an orthonormal set

of basis functions $\varphi_i : (\mathbf{r}, s) \mapsto \varphi_i(\mathbf{r}, s)$ in $H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$, that is,

$$\begin{aligned} B^d := \{\varphi_i : i = 1, \dots, d\} \quad &\subseteq \quad B := \{\varphi_i : i \in \mathbb{N}\} \quad \subseteq \quad H^1\Big(\mathbb{R}^3 \times \Big\{\pm\frac{1}{2}\Big\}\Big), \\ \mathcal{V}^d := \operatorname{Span} B^d \quad &\subseteq \quad \mathcal{V} := \operatorname{Span} B = H^1\Big(\mathbb{R}^3 \times \Big\{\pm\frac{1}{2}\Big\}\Big), \end{aligned}$$

where

$$\langle \varphi_i, \varphi_j \rangle := \sum_{s=\pm\frac{1}{2}} \int_{\mathbb{R}^3} \overline{\varphi_i(\mathbf{r}, s)} \varphi_j(\mathbf{r}, s) \mathrm{d}\mathbf{r} = \delta_{i,j}. \tag{9}$$

(For simplicity of notation, we will use the same brackets $\langle \cdot, \cdot \rangle$ for designating inner products in Hilbert spaces, independent of the underlying Hilbert space.) In quantum chemistry these functions are called *spin orbitals*, because they depend on the spin variable $s = \pm\frac{1}{2}$ and the spatial variable $\mathbf{r} \in \mathbb{R}^3$. In the sequel, we will first confine ourselves to spin orbital formulations. How we go from spin orbitals to *spatial orbitals* will be explained later.

We build *Slater determinants* of an N-electron system, by selecting N different indices, for example i_a for $a = 1, \dots, N$, out of the set $\{1, \dots, d\}$. By this we have chosen N orthonormal spin orbitals φ_{i_a}, $a = 1, \dots, N$, to define the *Slater determinant*[47,226]

$$\begin{aligned} \Phi_{[i_1,\dots,i_N]}(\mathbf{r}_1, s_1; \dots; \mathbf{r}_N, s_N) &= \frac{1}{\sqrt{N!}} \det\big(\varphi_{i_a}(\mathbf{r}_b, s_b)\big)_{a,b=1}^N \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} P(\sigma) \big(\varphi_{i_{\sigma(1)}} \otimes \cdots \otimes \varphi_{i_{\sigma(N)}}\big)(\mathbf{r}_1, s_1; \dots; \mathbf{r}_N, s_N), \end{aligned} \tag{10}$$

where the summation goes for all σ permutations of N elements, and $P(\sigma)$ is the parity of the permutation. To fix the sign of the determinant, we suppose e.g. that $i_a < i_{a+1}$ for $a = 1, \dots, N-1$; i.e., the indices are ordered increasingly. Therefore the Slater determinants are uniquely defined by referring to the orbital functions φ_{i_a}, respectively indices $i_a \in \{1, \dots, d\}$, which are contained in the determinant.

It is easy to check that the Slater determinants constructed in this way by the ortonormalized spin-orbitals $\varphi_i \in \mathcal{V}^d$ are also orthonormalized. We define the *Full Configuration Interaction (FCI) space for an N-electron system*[96,226] as the finite dimensional space $\mathcal{V}_N^d$ spanned by the Slater-determinants

$$\begin{aligned} B_N^d := \{\Phi_{[i_1,\dots,i_N]} : 1 \le i_a < i_{a+1} \le d\} \quad &\subseteq \quad B_N := \{\Phi_{[i_1,\dots,i_N]} : 1 \le i_a < i_{a+1}\} \quad \subseteq \quad \mathcal{V}_N, \\ \mathcal{V}_N^d := \operatorname{Span} B_N^d \quad &\subseteq \quad \mathcal{V}_N := \operatorname{Span} B_N. \end{aligned}$$

The dimension of $\mathcal{V}_N^d$ is

$$\dim \mathcal{V}_N^d = \binom{d}{N} = \frac{d!}{N!(d-N)!} \sim \mathcal{O}(d^N). \tag{11}$$

To obtain an approximate solution of the electronic Schrödinger equation, one may apply the *Ritz-Galerkin method* using the finite dimensional subspace $\mathcal{V}_N^d \subset \mathcal{V}_N$. I.e., consider the solution of the finite dimensional eigenvalue problem

$$\Pi_N^d H \Psi = E\Psi, \quad \Psi \in \mathcal{V}_N^d, \tag{12}$$

where $\Pi_N^d : \mathcal{V}_N \to \mathcal{V}_N^d$ is L_2-orthogonal projection, or, equivalently,

$$\langle \Phi, (H - E)\Psi \rangle = 0, \quad \Psi \in \mathcal{V}_N^d, \quad \text{for all } \Phi \in \mathcal{V}_N^d. \tag{13}$$

So the approximate ground state energy is

$$E_{0,d} := \min\{\langle \Psi, H\Psi \rangle : \langle \Psi, \Psi \rangle = 1, \Psi \in \mathcal{V}_N^d\}, \tag{14a}$$

and the *full CI* ground state wavefunction $\Psi_{0,d} \in \mathcal{V}_N^d$ is the solution of the Galerkin scheme

$$\Psi_{0,d} := \operatorname{argmin}\{\langle \Psi, H\Psi \rangle : \langle \Psi, \Psi \rangle = 1, \Psi \in \mathcal{V}_N^d\}. \tag{14b}$$

From the above definitions it becomes obvious that the approximate ground state energy $E_{0,d}$ obtained by the Ritz-Galerkin method provides an upper bound for the exact energy value $E_0 \leq E_{0,d}$, given in (7a). The convergence theory of the Galerkin scheme for the numerical solution of eigenvalue problems is well established[47,271]. Roughly speaking, the Galerkin method provides an approximate eigenfunction $\Psi_{0,d}$ which approximates the exact eigenfunction quasi-optimally. Moreover, the eigenvalue converges quadratically compared to the convergence of the eigenfunction. Since $\dim \mathcal{V}_d^N \sim \mathcal{O}(d^N) \geq \mathcal{O}(2^N)$, the full CI approach scales exponentially with respect to N. Therefore, for molecules this approach is practically not feasible, except for a very small number of electrons.

2.3 Fock spaces

We embed the full CI space $\mathcal{V}_N^d$ of N-electrons into a larger space $\mathcal{F}^d$, called *(discrete) Fock space*, where we do not care about the number of electrons,

$$\mathcal{F}^d := \bigoplus_{M=0}^{d} \mathcal{V}_M^d = \{\Psi = \Psi_{(0)} \oplus \Psi_{(1)} \oplus \ldots \oplus \Psi_{(d)} : \Psi_{(M)} \in \mathcal{V}_M^d\}. \tag{15}$$

Its dimension is

$$\dim \mathcal{F}^d = \sum_{M=0}^{d} \binom{d}{M} = 2^d. \tag{16}$$

The Fock space is again a Hilbert space with the inner product inherited from $\mathcal{V}_M^d$

$$\langle \bigoplus_{M=0}^{d} \Phi_{(M)}, \bigoplus_{M'=0}^{d} \Psi_{(M')} \rangle = \sum_{M=0}^{d} \langle \Phi_{(M)}, \Psi_{(M)} \rangle. \tag{17}$$

The full Fock space can be obtained by taking the limit for $d \to \infty$. Since we consider only finite dimensional approximation, we are not intended here to understand in what sense this limit might be defined or not.

We have the Hamiltonian (5a) acting on the full-CI space of N electrons as $\Pi_N^d H : \mathcal{V}_N^d \to \mathcal{V}_N^d$. Since now we allow different numbers of electrons, we denote this explicitly as H_N, then the Hamiltonian acting on the whole Fock space reads

$$H^d := \bigoplus_{M=0}^{d} \Pi_M^d H_M \ : \ \mathcal{F}^d \to \mathcal{F}^d. \tag{18}$$

It is convenient to define the *creation operator* $a_i^\dagger : \mathcal{F}^d \to \mathcal{F}^d$, which is given on Slater determinants as $a_i^\dagger \Phi_{[i_1,\dots,i_N]} := \Phi_{[i,i_1,\dots,i_N]}$. This connects the subspaces with different numbers of particles in the Fock space, $a_i^\dagger : \mathcal{V}_N^d \to \mathcal{V}_{N+1}^d$. The result of this operator acting on a Slater determinant in B_N^d is a Slater determinant again, and, up to a $\pm$ sign, it is contained in B_N^d:

$$a_i^\dagger \Phi_{[i_1,\dots,i_N]} = \Phi_{[i,i_1,\dots,i_N]} = (-1)^{k_i} \Phi_{[i_1,\dots,i,\dots,i_N]}, \tag{19a}$$

where the indices in $[i_1,\dots,i,\dots i_N]$ are ordered increasingly, and $k_i = |\{i_b | i_b < i\}|$. From the definition, it immediately follows that $a_i^\dagger \Phi_{[i_1,\dots,i_N]} = 0$ if $i \in [i_1,\dots,i_N]$, which is the manifestation of the exclusion principle. One can then obtain the adjoint $a_i := (a_i^\dagger)^\dagger$ of the creation operator, which is called *annihilation operator*

$$a_i \Phi_{[i_1,\dots,i_b,\dots,i_N]} = (-1)^{k_i} \Phi_{[i_1,\dots,i_N]}, \qquad \text{if } i_b = i \text{ for some } b = 1,\dots,N, \text{ othervise } 0. \tag{19b}$$

It is straightforward to check that these operators obey the *fermionic anticommutation relations:*

$$\{a_i^\dagger, a_j^\dagger\} = 0, \qquad \{a_i, a_j\} = 0, \qquad \{a_i, a_j^\dagger\} = \delta_{i,j}, \tag{20}$$

(with the *anticommutator* $\{A, B\} = AB + BA$), which is the manifestation of the Pauli antisymmetry principle (3). One can check that the operator $n_i := a_i^\dagger a_i$ leaves invariant all the Slater determinants for which $i \in [i_1,\dots,i_N]$, while it annihilates all the Slater determinants for which $i \notin [i_1,\dots,i_N]$. One can conclude then that the operator $P = \sum_{i=1}^d a_i^\dagger a_i$ acts on $\mathcal{V}_M^d$ as M times the identity, that is

$$P = \sum_{i=1}^d a_i^\dagger a_i = \bigoplus_{M=0}^d M I_{\mathcal{V}_N^d} \tag{21}$$

on the whole Fock space $\mathcal{F}^d$, and it is called *particle number operator*, since

$$P(0 \oplus \cdots \oplus \Psi_{(M)} \oplus \cdots \oplus 0) = M\ 0 \oplus \cdots \oplus \Psi_{(M)} \oplus \cdots \oplus 0.$$

2.4 Occupation numbers and second quantization

Instead of the above notations, it is usual to introduce the very convenient binary labeling, or *occupation number labeling*, for the Slater determinants $\Phi_{[i_1,\dots,i_N]}$. Let $(\mu_1,\dots,\mu_d)$ be a binary string, i.e., $\mu_i \in \{0,1\}$, depending on the presence or absence of φ_i in the Slater determinant $\Phi_{[i_1,\dots,i_N]}$: For all $i = 1,\dots,d$, if $i \in [i_1,\dots,i_N]$ then $\mu_i = 1$ (then we say that spin-orbital φ_i is *occupied* in the Slater determinant), else $\mu_i = 0$ (*unoccupied*). So $\mu_i \in \{0,1\}$ has the meaning of an *occupation number*, and we use the notation

$$\Phi_{(\mu_1,\dots,\mu_d)} := \Phi_{[i_1,\dots,i_N]}, \qquad \mu_i \in \{0,1\}, i = 1,\dots,d. \tag{22}$$

Furthermore, in an N particle Slater determinant, $\mu_i = 1$ appears exactly N times. With this, the Fock space becomes

$$\mathcal{F}^d = \bigoplus_{M=0}^d \mathcal{V}_M^d = \Big\{ \Psi : \Psi = \sum_{\mu_1,\dots,\mu_d} u_{\mu_1,\dots,\mu_d} \Phi_{(\mu_1,\dots,\mu_d)}, u_{\mu_1,\dots,\mu_d} \in \mathbb{C}, \mu_i \in \{0,1\}, i = 1,\dots,d \Big\}. \tag{23}$$

The effect of the creation and annihilation operators (19) can also be formulated using the occupation numbers in a more expressive way than before:

$$a_i^\dagger \Phi_{(\mu_1,\dots,\mu_i,\dots,\mu_d)} = (-1)^{k_i} \Phi_{(\mu_1,\dots,\mu_i+1,\dots,\mu_d)}, \quad \text{if } \mu_i = 0, \text{ otherwise } 0, \tag{24a}$$

$$a_i \Phi_{(\mu_1,\dots,\mu_i,\dots,\mu_d)} = (-1)^{k_i} \Phi_{(\mu_1,\dots,\mu_i-1,\dots,\mu_d)}, \quad \text{if } \mu_i = 1, \text{ otherwise } 0, \tag{24b}$$

where $k_i = \sum_{j=1}^{i-1} \mu_j$. On the other hand, $a_i^\dagger a_i \Phi_{(\mu_1,\dots,\mu_i,\dots,\mu_d)} = \mu_i \Phi_{(\mu_1,\dots,\mu_i,\dots,\mu_d)}$, and with the definition (21) we have

$$P\Phi_{(\mu_1,\dots,\mu_d)} = \Big(\sum_{i=1}^{d} \mu_i\Big) \Phi_{(\mu_1,\dots,\mu_d)}. \tag{25}$$

This binary labelling gives us the opportunity of using another, more convenient, representation of the Fock space, which is called *second quantization*. To this end, we consider the Hilbert space for the representation of the events of the occupation of the orbitals. This space has a two dimensional tensor factor $\Lambda_i \cong \mathbb{C}^2$ for each orbital, containing two ortogonal states representing the unoccupied and occupied states of the orbital. So, let us define the tensor product space

$$\Lambda^{(d)} := \bigotimes_{i=1}^{d} \Lambda_i \cong \bigotimes_{i=1}^{d} \mathbb{C}^2, \tag{26}$$

and the canonical basis $\{|\phi_0\rangle \cong \mathbf{e}_0, |\phi_1\rangle \cong \mathbf{e}_1\}$ of the vector space $\Lambda_i \cong \mathbb{C}^2$, where $(\mathbf{e}_0)_\mu = \delta_{\mu,0}$, $(\mathbf{e}_1)_\mu = \delta_{\mu,1}$. (For the elements of only these spaces we use the so called *ket*-notation $|\dots\rangle$, which is very common in physics for denoting elements of Hilbert spaces.) So we write any $|U\rangle \in \mathbb{C}^2$ as $|U\rangle = \sum_{\mu=0}^{1} U(\mu)|\phi_\mu\rangle$. The dimension of this space is

$$\dim \Lambda^{(d)} = 2^d. \tag{27}$$

We also have the canonical inner product in $\mathbb{C}^2$:

$$\langle U|V\rangle = \sum_{\mu=0}^{1} \overline{U(\mu)} V(\mu), \tag{28}$$

for which the canonical basis is orthogonal, $\langle \phi_\mu | \phi_\nu \rangle = \delta_{\mu,\nu}$, for $\mu, \nu = 0, 1$. Using the canonical basis $\{|\phi_{\mu_i}^{\{i\}}\rangle\}$ of Λ_i, the set $\{|\phi_{\mu_1}^{\{1\}}\rangle \otimes \cdots \otimes |\phi_{\mu_d}^{\{d\}}\rangle : \mu_i \in \{0,1\}, i = 1,\dots,d\}$ gives the canonical basis in $\Lambda^{(d)}$, and any $|U\rangle \in \Lambda^{(d)}$ can be represented by

$$|U\rangle = \sum_{\mu_1,\dots,\mu_d} U(\mu_1,\dots,\mu_d) |\phi_{\mu_1}^{\{1\}}\rangle \otimes \cdots \otimes |\phi_{\mu_d}^{\{d\}}\rangle. \tag{29}$$

If there is no ambiguity about the underlying basis, we consider the canonical tensor product basis above, and we can identify $|U\rangle \in \Lambda^{(d)}$ with $U \in \bigoplus_{i=1}^{d} \mathbb{C}^2$, where U are simply d-variate functions (see section 3.1)

$$(\mu_1,\dots,\mu_d) \quad \longmapsto \quad U(\mu_1,\dots,\mu_d) \in \mathbb{C}, \tag{30}$$

depending on the discrete variables $\mu_i = 0, 1$, for $i = 1, \ldots, d$, called also indices in the sequel. Due to the orthogonality of the canonical basis, the canonical inner product of $\mathbb{C}^2$ induces

$$\langle U|V\rangle := \sum_{\mu_1,\ldots,\mu_d} \overline{U(\mu_1,\ldots,\mu_d)} V(\mu_1,\ldots,\mu_d) \tag{31}$$

and the norm $\|U\| = \sqrt{\langle U, U\rangle}$ in $\Lambda^{(d)}$.

Now, let us introduce the *isomorphism* $\iota : \mathcal{F}^d \to \Lambda^{(d)}$, defined by its action on the basis functions, i.e., the Slater determinants $\Phi_{(\mu_1,\ldots,\mu_d)}$ of occupation numbers $(\mu_1, \ldots, \mu_d)$ simply as

$$\iota(\Phi_{(\mu_1,\ldots,\mu_d)}) := |\phi_{\mu_1}^{\{1\}}\rangle \otimes \cdots \otimes |\phi_{\mu_d}^{\{d\}}\rangle, \qquad \mu_i \in \{0,1\}, \quad i = 1, \ldots, d. \tag{32}$$

(This is an elementary tensor product, in physics called *tensor product state*.) It is easy to check that this is indeed an isomorphism, and compatible with the inner products of the two spaces, so we conclude that the discrete Fock space $\mathcal{F}^d$ is isomorphic to $\Lambda^{(d)} \cong \otimes_{i=1}^d \mathbb{C}^2$, and

$$|\Psi\rangle := \iota(\Psi) = \iota\Big(\sum_{\mu_1,\ldots,\mu_d} u_{\mu_1,\ldots,\mu_d} \Phi_{(\mu_1,\ldots,\mu_d)}\Big) = \sum_{\mu} U(\mu_1,\ldots,\mu_d) |\phi_{\mu_1}^{\{1\}}\rangle \otimes \cdots \otimes |\phi_{\mu_d}^{\{d\}}\rangle \tag{33a}$$

leads to

$$U(\mu_1,\ldots,\mu_d) = u_{\mu_1,\ldots,\mu_d}. \tag{33b}$$

On the other hand, we have used the convention above that for a function $\Psi \in \mathcal{F}^d$, its image by ι is written as the *ket* $|\Psi\rangle$. The Full-CI space for N electrons $\mathcal{V}_N^d$ is a subspace of the Fock space $\mathcal{F}^d$, and its image in $\Lambda^{(d)}$ is denoted as $\Lambda_{\mathrm{FCI}} := \iota(\mathcal{V}_N^d) \subset \Lambda^{(d)}$. This is the N-electron subspace of $\Lambda^{(d)}$, having dimension $\dim \Lambda_{\mathrm{FCI}} = \binom{d}{N}$.

Through this isomorphism, we can obtain the creation and annihilation operators (24) acting on $\Lambda^{(d)}$ as follows[197]

$$\mathbf{a}_i^\dagger := \iota \circ a_i^\dagger \circ \iota^{-1} = \mathbf{s} \otimes \ldots \otimes \mathbf{s} \otimes \mathbf{a}^\dagger \otimes \mathbf{I} \otimes \ldots \otimes \mathbf{I} \ : \ \Lambda^{(d)} \to \Lambda^{(d)}, \tag{34a}$$

$$\mathbf{a}_i := \iota \circ a_i \circ \iota^{-1} = \mathbf{s} \otimes \ldots \otimes \mathbf{s} \otimes \mathbf{a} \otimes \mathbf{I} \otimes \ldots \otimes \mathbf{I} \ : \ \Lambda^{(d)} \to \Lambda^{(d)}, \tag{34b}$$

(where $a^\dagger$ and a appear in the i-th position) with the operators

$$\mathbf{a} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{a}^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{s} := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{I} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{35}$$

acting on Λ_i. The operators $\mathbf{a}_i$ and $\mathbf{a}_i^\dagger$ again obey the fermionic anticommutation relation (20). Let us highlight that the 2×2-matrix $\mathbf{s}$ is required to provide the correct phase factor, i.e., sign of the Slater determinant. The particle number operator acting on $\Lambda^{(d)}$ is

$$\mathbf{P} := \iota \circ P \circ \iota^{-1} = \sum_{i=1}^d \mathbf{a}_i^\dagger \mathbf{a}_i, \tag{36}$$

which is, since $\mathbf{s}^2 = I$, the sum of matrices

$$\mathbf{n}_i = \mathbf{a}_i^\dagger \mathbf{a}_i = \mathbf{I} \otimes \ldots \otimes \mathbf{I} \otimes \mathbf{a}^\dagger \mathbf{a} \otimes \mathbf{I} \otimes \ldots \otimes \mathbf{I} \ : \ \Lambda^{(d)} \to \Lambda^{(d)}, \tag{37}$$

with the matrix

$$\mathbf{n} = \mathbf{a}^\dagger \mathbf{a} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{38}$$

in the i-th position, representing the occupation number of the given orbital.

Let us remark that since the isomorphism ι is defined through a given Slater determinant basis, the above representation of a wave function is basis dependent, i.e., it depends on the choice of the one-particle basis set B^d.

2.5 Example: Hartree-Fock determinant and change of one-particle basis

The Hartree Fock determinant in terms of canonical molecular orbital functions is given by $\Psi_{\mathrm{HF}} = \Phi_{[1,\dots,N]} = \Phi_{(1,\dots,1,0,\dots,0)}$, i.e., $i_b = b$ for $b = 1, \dots, N$, the *first* N spin-orbitals are occupied. In $\Lambda^{(d)}$ this is represented by the tensor for which $U(\mu_1, \dots, \mu_d) = 1$ if and only if $(\mu_1, \dots, \mu_d) = (1, \dots, 1, 0, \dots, 0)$, or

$$|\Psi_{\mathrm{HF}}\rangle = |\phi_1^{\{1\}}\rangle \otimes \cdots \otimes |\phi_1^{\{N\}}\rangle \otimes |\phi_0^{\{N+1\}}\rangle \otimes \cdots \otimes |\phi_0^{\{d\}}\rangle. \tag{39}$$

If we move to another basis, respectively basis set, say

$$B^d = \big\{\varphi_i : i = 1, \dots, d\big\} \qquad \longmapsto \qquad \widetilde{B^d} = \Big\{\widetilde{\varphi}_i = \sum_{j=1}^{d} U_{i,j}\varphi_j : i = 1, \dots, d\Big\}, \tag{40}$$

with the unitary $d \times d$ matrix $\mathbf{U} = \big(U_{i,j}\big)$, the representation of the old Slater determinants $\Phi_{(\mu_1,\dots,\mu_d)} = \sum_{\nu_1,\dots,\nu_d} \widetilde{v}_{\nu_1,\dots,\nu_d} \widetilde{\Phi}_{(\nu_1,\dots,\nu_d)}$ in terms of the new Slater determinants $\widetilde{\Phi}_{(\nu_1,\dots,\nu_d)}$ is no longer of rank one. It is a short exercise to show a representation of the form $\Psi = \sum_{\mu_1,\dots,\mu_d} u_{\mu_1,\dots,\mu_d} \Phi_{(\mu_1,\dots,\mu_d)}$ is transformed into

$$\begin{aligned} |\widetilde{\Psi}\rangle &= e^{\sum_{i=N+1}^{d} \sum_{j=1}^{N} t_{i,j} \mathbf{a}_i^\dagger \mathbf{a}_j} |\Psi\rangle \\ &= \Pi_{i=N+1}^{d} \Pi_{j=1}^{N} (\mathbf{I}^d + t_{i,j} \mathbf{a}_i^\dagger \mathbf{a}_j)\, |\Psi\rangle \\ &= (\mathbf{I}^d + t_{N+1,1} \mathbf{a}_{N+1}^\dagger \mathbf{a}_1) \cdot \ldots \cdot (\mathbf{I}^d + t_{d,N} \mathbf{a}_d^\dagger \mathbf{a}_N) |\Psi\rangle \end{aligned}$$

up to a normalization constant. This transformation serves as the transformation of the basis sets. Let us remark that in Coupled Cluster theory the above expression is known as the *Coupled Cluster single excited states*, and the above transformation is used for the definition of the *Brückner orbital*[96]. Also each factor is at most of rank two, so in the worst case the rank could increased by a factor of two with each matrix multiplication. The single Slater determinant expressed by another basis set is a linear combination of many Slater determinants with respect to the new basis. Indeed it can happen that in the new bases it is represented by the maximally entangled state tensor $|\widetilde{\Psi}\rangle$.

2.6 Ritz-Galerkin approximation in second quantization

Now we are able to formulate the full-CI Schrodinger equation (12) in the second-quantized form. By the isomorphism ι, the Hamiltonian (18) acting on $\mathcal{F}^d$ is defined on $\Lambda^{(d)}$ as follows

$$\mathbf{H} = \iota \circ H^d \circ \iota^{-1} \ : \ \Lambda^{(d)} \to \Lambda^{(d)}. \tag{41}$$

Using the ansatz $\Psi = \sum_{\nu_1,\dots,\nu_d} u_{\nu_1,\dots,\nu_d}\Phi_{(\nu_1,\dots,\nu_d)} \in \mathcal{V}_N^d$ for the eigenvector, with the help of ι, we obtain the discrete eigenvalue problem from (13) in the N-electron subspace of $\Lambda^{(d)}$

$$\begin{aligned}\langle \Phi_{(\mu_1,\dots,\mu_d)}, (H-E)\Psi\rangle &= \sum_{\nu_1,\dots,\nu_d} \langle \Phi_{(\mu_1,\dots,\mu_d)}, (H-E)\Phi_{(\nu_1,\dots,\nu_d)}\rangle u_{\nu_1,\dots,\nu_d} \\ &= \big((\mathbf{H}-E\mathbf{I})|\Psi\rangle\big)_{\mu_1,\dots,\mu_d} = 0 \quad \forall \mu_i \in \{0,1\}, \quad i=1,\dots,d.\end{aligned}$$

If we allow $|\Psi\rangle \in \Lambda^{(d)}$, not only in its N-electron subspace, we do not get only the eigenvalues E for a discretized N electron system, but also the eigenvalues for other numbers of electrons between 1 and d. To fix this kind of problem, we take into account the particle number operator (36), then $\Psi = \sum_{\mu_1,\dots,\mu_d} u_{\mu_1,\dots,\mu_d}\Phi_{(\mu_1,\dots,\mu_d)} \in \mathcal{V}_N^d$ holds if and only if

$$\mathbf{P}|\Psi\rangle = N|\Psi\rangle, \tag{42}$$

i.e., $|\Psi\rangle$ is an eigenvector of $\mathbf{P}$ with the eigenvalue N.

From the well known Slater-Condon rules[226] one concludes the precise representation of the Hamiltonian in the discrete space $\Lambda^{(d)}$, which reads as

$$\mathbf{H} = \sum_{i,j=1}^{d} T_{ij}\mathbf{a}_i^\dagger \mathbf{a}_j + \sum_{i,j,k,l=1}^{d} V_{ijkl}\mathbf{a}_i^\dagger \mathbf{a}_j^\dagger \mathbf{a}_k \mathbf{a}_l. \tag{43}$$

Here the coefficients

$$T_{ij} = \sum_{s=\pm\frac{1}{2}} \int_{\mathbb{R}^3} \overline{\varphi_i(\mathbf{r},s)}\Big(-\frac{1}{2}\Delta - \sum_{c=1}^{K}\frac{Z_c}{|\mathbf{r}-\mathbf{R}_c|}\Big)\varphi_j(\mathbf{r},s)\mathrm{d}\mathbf{r} \tag{44a}$$

are the well known *single electron integrals,* and

$$V_{ijkl} = \sum_{s,s'=\pm\frac{1}{2}} \int_{\mathbb{R}^3} \overline{\varphi_i(\mathbf{r},s)\varphi_j(\mathbf{r}',s')}\frac{1}{|\mathbf{r}'-\mathbf{r}|}\varphi_k(\mathbf{r},s)\varphi_l(\mathbf{r}',s')\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \tag{44b}$$

are the *two electron integrals,* both are coming from the parts of the original Hamiltonian (5b). With this, the discrete (Full CI) Schrödinger equation for the approximation of the ground state can be cast into the binary variational form of finding $|\Psi\rangle \in \Lambda^{(d)}$ such that

$$E_{0,d} = \min\big\{\langle\Psi|\mathbf{H}|\Psi\rangle : \langle\Psi|\Psi\rangle = 1,\ \mathbf{P}|\Psi\rangle = N|\Psi\rangle,\ |\Psi\rangle \in \Lambda^{(d)}\big\}, \tag{45a}$$

$$|\Psi_{0,d}\rangle = \operatorname{argmin}\big\{\langle\Psi|\mathbf{H}|\Psi\rangle : \langle\Psi|\Psi\rangle = 1,\ \mathbf{P}|\Psi\rangle = N|\Psi\rangle,\ |\Psi\rangle \in \Lambda^{(d)}\big\}. \tag{45b}$$

Let us remark that this representation depends on the basis orbital functions B^d. For a change of basis, as is given in (40), the creation and annihilation operators transform as

$$\widetilde{\mathbf{a}}_i^\dagger = \sum_{j=1}^{d} U_{i,j}\mathbf{a}_j^\dagger, \qquad \widetilde{\mathbf{a}}_i = \sum_{j=1}^{d}\overline{U_{i,j}}\mathbf{a}_j. \tag{46}$$

With respect to the new basis set $\widetilde{B^d}$, we can build another Slater determinant basis $\{\widetilde{\Phi}_{(\nu_1,\dots,\nu_d)} : \nu_i \in \{0,1\}, i = 1,\dots,d\}$ of the discrete Fock space $\mathcal{F}^d$. With respect to this new Slater determinant basis, the operators $\widetilde{\mathbf{a}}_i$ has the canonical form (34b)

$$\widetilde{\mathbf{a}}_i = \mathbf{s} \otimes \cdots \otimes \mathbf{s} \otimes \mathbf{a} \otimes \mathbf{I} \otimes \cdots \otimes \mathbf{I}. \tag{47}$$

The one and two electron integrals transform easily, e.g.,

$$\tilde{T}_{ij} = \sum_{kl} (U^\dagger)_{ik} T_{kl} U_{lj}. \tag{48}$$

2.7 Spatial orbitals

For the sake of simplicity of representation, throughout this section we have dealt with spin orbital basis φ_i for $i = 1,\dots,d$, and the above setting $\Lambda^{(d)} = \bigotimes_{i=1}^d \Lambda_i \cong \bigotimes_{i=1}^d \mathbb{C}^2$. In the QC-DMRG, it has been experienced that it is favorable to use spin functions explicitly, and deal only with *spatial orbitals*. More precisely, we take a set $\{\kappa_i \in H^1(\mathbb{R}^3) : i = 1,\dots,d\}$ of orthonormalized functions, depending on the space-variable $\mathbf{r} \in \mathbb{R}^3$ only, and define the basis of $2d$ elements

$$\varphi_{2i}(\mathbf{r}, s) = \kappa_i(\mathbf{r})\chi_+(s), \qquad \chi_+(s) = 1 \quad \text{if } s = +\frac{1}{2}, \qquad \chi_+(s) = 0 \quad \text{if } s = -\frac{1}{2}, \tag{49a}$$

$$\varphi_{2i+1}(\mathbf{r}, s) = \kappa_i(\mathbf{r})\chi_-(s), \qquad \chi_-(s) = 0 \quad \text{if } s = +\frac{1}{2}, \qquad \chi_-(s) = 1 \quad \text{if } s = -\frac{1}{2}, \tag{49b}$$

which are orthonormalized in $H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$. Now, repeating the previous construction leads to the $2d$ spaces $W_{2i-1} \cong \mathbb{C}^2$ and $W_{2i} \cong \mathbb{C}^2$. Let us cast the tensor product of two adjacent spaces into one $\Lambda_i := W_{2i-1} \otimes W_{2i} \cong \mathbb{C}^4$, and with this,

$$\Lambda^{(d)} = \bigotimes_{i=1}^d \Lambda_i \cong \bigotimes_{i=1}^d \mathbb{C}^4, \tag{50}$$

having the dimension $\dim \Lambda^{(d)} = 4^d$. The N-electron subspace Λ_{FCI} is then of dimension $\dim \Lambda_{\mathrm{FCI}} = \binom{2d}{N}$. In the case when the $N_\downarrow$ and $N_\uparrow$ numbers of electrons of spins $-1/2$ and $+1/2$ are conserved, only a subspace of this is needed, which is called then the Full-CI space that is of dimension $\dim \Lambda_{\mathrm{FCI}} = \binom{d}{N_\downarrow}\binom{d}{N_\uparrow} < \binom{2d}{N}$.

Using the matrices (35) we define

$$\mathbf{c}_s := \mathbf{a} \otimes \mathbf{I} \quad \text{if } s = +\frac{1}{2}, \qquad \mathbf{c}_s := \mathbf{s} \otimes \mathbf{a} \quad \text{if } s = -\frac{1}{2}, \tag{51a}$$

$$\mathbf{z} := \mathbf{s} \otimes \mathbf{s}, \qquad \mathbb{I} = \mathbf{I} \otimes \mathbf{I}, \tag{51b}$$

$$\mathbf{c}_{i,s} := \mathbf{z} \otimes \cdots \otimes \mathbf{z} \otimes \mathbf{c}_s \otimes \mathbb{I} \otimes \cdots \otimes \mathbb{I}, \tag{51c}$$

$$\mathbf{n}_s := \mathbf{c}_s^\dagger \mathbf{c}_s, \qquad \mathbf{n}_{i,s} := \mathbf{c}_{i,s}^\dagger \mathbf{c}_{i,s} = \mathbb{I} \otimes \cdots \otimes \mathbb{I} \otimes \mathbf{c}_s^\dagger \mathbf{c}_s \otimes \mathbb{I} \otimes \cdots \otimes \mathbb{I}. \tag{51d}$$

With these, the Hamilton operator reads as

$$\mathbf{H} = \sum_{i,j=1}^d \sum_{s_i,s_j=\pm\frac{1}{2}} T_{ij} \mathbf{c}_{i,s_i}^\dagger \mathbf{c}_{j,s_j} + \sum_{i,j,k,l=1}^d \sum_{s_i,s_j,s_k,s_l=\pm\frac{1}{2}} V_{ijkl} \mathbf{c}_{i,s_i}^\dagger \mathbf{c}_{j,s_j}^\dagger \mathbf{c}_{k,s_k} \mathbf{c}_{l,s_l}. \tag{52}$$

Let us remark that in the nonrelativistic quantum chemistry, the one- and two electron integrals do not depend on the spin variables s_i. This is the reason for using the spatial orbital formulation. (However, this does not hold when relativistic effects, e.g. spin-orbit coupling are taken into account, as is used in a recent development in DMRG[109].)

3 Tensor product approximation

3.1 Tensor product parametrization

We generalize (33a) by considering vector spaces with arbitrary dimension $\dim \Lambda_i = q_i$:

$$|\Psi\rangle = \sum_{\alpha_1,\ldots,\alpha_d} U(\alpha_1,\ldots,\alpha_d)\ |\phi^{\{1\}}_{\alpha_1}\rangle \otimes \cdots \otimes |\phi^{\{d\}}_{\alpha_d}\rangle. \tag{53}$$

Thus, the tensor

$$|\Psi\rangle \in \Lambda^{(d)} \tag{54}$$

is equivalent to the multi-indexed array

$$U \in \mathbb{C}^{q_1\times\ldots\times q_d}, \tag{55}$$

$$U(\alpha_1,\ldots,\alpha_d),\ \alpha_i \in \{1,\ldots,q_i\}\quad i = 1,\ldots,d. \tag{56}$$

The norm and inner product of these two tensor spaces are defined analogously to Sec. 2.4.

Computation with tensors suffer from the *curse of dimensionality*[83], since the storage of the complete array grows exponentially with the order d.

We seek to reduce computational costs by parametrizing the tensors in some data-sparse representation. For this purpose, we adhere to the *separation of variables*, a classical approach which traces back to Bernoulli and Fourier among others. In principle, we want to represent or approximate tensors as multi-variate functions by a sum of products of univariate functions. This concept is well established for tensors of order $d = 2$ where it leads to fundamental results known as the *singular value decomposition* (SVD) or Schmidt decomposition, proper orthogonal decomposition, the Karhunen-Loeve transform and so on. In the discrete case discussed here, i.e. in matrix theory, this is known as *low rank approximation*. However, the generalization of the concept of ranks to higher order tensors is not as straightforward as one may expect[82]. There are many possible and a priori equally justifiable tensor decompositions that all yield different definitions of a tensor rank.

The *canonical tensor representation* separates the variables

$$U(\alpha_1,\ldots,\alpha_d) = \sum_{i=1}^{R} u_i^1(\alpha_1)\cdots u_i^d(\alpha_d). \tag{57}$$

The canonical tensor rank of U is the smallest R such that this representation is exact. This is then called the *canonical decomposition* of the tensor[82].

However, while this is a beautiful and rather simplistic tensor representation, it has several severe drawbacks. First of all, finding the canonical rank and thus also its decomposition is NP-hard[110]. Additionally, the set of tensors with rank smaller or equal R is not closed,

i.e. it is possible to find a sequence of rank-R-tensors that converges to a tensor with rank greater than R, see the *border rank problem*[82,119]. While the former property obviously poses problems in computation, the latter can be very undesirable as well when it comes to optimization algorithms. Altogether, the canonical format has not only led to deep and difficult mathematical problems[119,120], but also computational experience has often been disappointing, by observing slow convergence, low accuracy and the like. It is not clear how to circumvent these problems while still retaining its outstanding complexity scaling. In recent years, the canonical format has therefore been put into question, albeit not completely disqualified, and we are looking for alternatives with favorable properties.

We parametrize a tensor in a very general form to define a *tensor representation* via

$$U(\alpha_1, \dots, \alpha_d) = \sum_{\mathbf{m}}^{\mathbf{r}} \prod_{i}^{K} U_i(\alpha_{i,1}, \dots, \alpha_{i,y_i}; m_{i,1}, \dots, m_{i,z_i}), \tag{58}$$

where $\mathbf{m}$ denotes the multi-index

$$\mathbf{m} = \bigcup_{i=1}^{K} \{m_{i,1}, \dots, m_{i,z_i}\}. \tag{59}$$

Since the ordering of the indices is irrelevant in this context, we maintain the slight abuse of notation and interpret multi-indices as sets of natural numbers.

This tensor representation is parametrized by K *component tensors* $U_1, \dots, U_K$. The α_i are called *physical* indices and the m_j are called *virtual*. A component U_i is called virtual, if it does not have any physical indeces, i.e. $y_i = 0$. Otherwise it is called physical. Summation over a common virtual index m_j is called the *contraction* over m_j.

We can demand a number of further properties that allow for simpler treatment of the tensor. First of all, it is conventional to only deal with multi-linear representations:

Criterion 3.1. For each $j \in \{1, \dots, d\}$ there exists *exactly* one $i \in \{1, \dots, K\}$ such that $\alpha_j \in \{\alpha_{i,1}, \dots, \alpha_{i,y_i}\}$.

This means that no two components can depend on the same physical index. The present multi-linear parametrization provides simple representation of the derivatives, an indispensable tool for local optimization[65,66], and alternating directional search methods[99].

It is our central aim to reduce complexity of the tensor and we therefore need to choose the representation carefully. Firstly, the number of components should not be exceedingly high, as this makes the representation more complicated. But more importantly, we try to minimize the dimensions $\mathbf{r}$ of the multi-index $\mathbf{m}$ over all possible representations (58). If these dimensions r_i are minimal for the given parametrization, the tuple $\mathbf{r}$ is called the *rank*, or better *multi-linear rank* of the representation and the representation is called a *decomposition*. However, as mentioned for the canonical format above, this notion of rank leads to extreme difficulties even for the simplest forms.

3.2 Tensor networks

For a proper definition of multi-linear ranks, we consider subclasses of tensor representations and introduce a further restriction that each $m_i \in \mathbf{m}$ appears exactly twice:

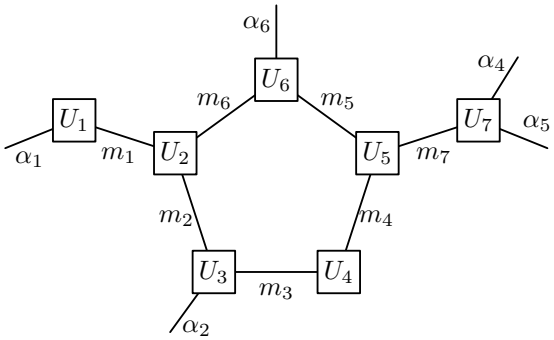


Figure 1: A general tensor network representation of a tensor of order 5.

Criterion 3.2. For each virtual index $m_i \in \mathbf{m}$ there exist *exactly* two component tensors U_{i_1}, U_{i_2} with m_i as an index.

Any parametrization satisfying criterion 3.1 and 3.2 can be expressed as a simple undirected weighted graph with half-edges, and we obtain what is called a *tensor network* or *tensor network states* in quantum physics. The component tensors give the vertices of the graph, the contractions are represented by the edges between the vertices and the physical indices yield half edges. Therefore, we get a graph $TNS(U) := (V, E, H)$,

$$V = \{U_i : i = 1, \dots, K\},\ E = \mathbf{m},\ H = \{\alpha_1, \dots, \alpha_d\}. \tag{60}$$

Because of criterion 3.1, each half-edge has exactly one incident vertex, and because of 3.2, each edge has exactly two incident vertices[65]. Thus, this is well-defined. The weight of the half-edge α_i is given by its dimension q_i and the weight of the edge m_j is given by its dimension r_j. In accordance with the tensor decompositions, we call the vector $\mathbf{r}$ the *rank* of the tensor network if it is minimal. $q_1 \cdots q_d$ is naturally the dimension of the tensor network, as shown in Fig. 1.

Since a contraction over an index with dimension 1 is trivial, we can choose to either omit this index or even to introduce extra indices. In general, we require the tensor network graph to be connected and if it is not we invent an arbitrary index of dimension 1 to make it so. Apart from that, any index of dimension 1 that is not necessary for connectedness will usually be omitted.

Although heavily used in physics this general concept still suffers from some instabilities. Recently it has been shown that tensor networks which contain closed loops are not necessarily Zariski closed[120], i.e. they do not form algebraic varieties without further restrictions. This is closely related to the border rank problem for the canonical format. While we will not go into these details here, we highlight that all these difficulties can be avoided, if we restrict ourselves to tensors fulfilling the following criterion[120]:

Criterion 3.3. The tensor network $TNS(U)$ is cycle-free.

Since we have the trivial connectedness mentioned above, any tensor network that fulfills criterion 3.3 is a tree. It is thus called a *Tree Tensor Network* or, in accordance with nomenclature from Quantum Physics, *Tree Tensor Network States (TTNS)*. See Fig. 2 for an arbitrary example.

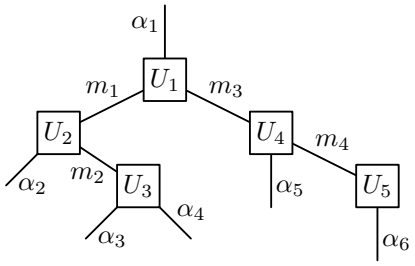


Figure 2: An arbitrary example of a tensor tree.

While general tensor network representations, like the canonical format, might still be very useful and shall not be disqualified, we presently only consider the special case of non-circular graphs that prevents these fundamental difficulties.

3.3 Subspace optimization and the Tucker format

Tensor trees can be introduced from a different perspective in order to illustrate that they share favorable properties with the matrix case. In our discrete setting, the tensor space and the tensor product space are equivalent

$$\mathbb{C}^{q_1 \times \ldots \times q_d} \cong \bigotimes_{i=1}^{d} \mathbb{C}^{q_i} \cong \bigotimes_{i=1}^{d} \Lambda_i \tag{61}$$

via the trivial formula (53). $|\phi_{\alpha_i}^{\{i\}}\rangle$ are the standard euclidean basis vectors of Λ_i for each i.

In this context, we define the *subspace optimization* as the best approximation of U[82,84]

$$\operatorname{argmin}\{\|U - U_\epsilon\| : U_\epsilon \in \bigotimes_{i=1}^{d} \Xi_i,\ \dim \Xi_i \leq r_i\}, \tag{62}$$

where we optimize over the tensor product of all univariate subspaces $\Xi_i \subset \Lambda_i$ of dimension at most r_i. If we can recover the tensor U exactly, i.e. $\|U - U_\epsilon\| = 0$, we call U_ϵ the *subspace representation* of U. In accordance with the above, a subspace representation is called a decomposition if the dimensions r_i are the ranks, i.e. they are the smallest numbers such that the tensor can still be recovered exactly.

This immediately motivates the *Tucker decomposition format* of a tensor. For each $i = 1, \ldots, d$, we aim at finding an optimal basis set $\{|\xi_{m_i}^{\{i\}}\rangle : m_i = 1, \ldots, r_i\}$ of a subspace $\Xi_i \subseteq \mathbb{C}^{q_i}$ where $r_i \leq q_i$. (53) can thus be restated as

$$|\Psi\rangle = \sum_{m_1=1}^{r_1} \ldots \sum_{m_d=1}^{r_d} C(m_1, \ldots, m_d)\ |\xi_{m_1}^{\{1\}}\rangle \otimes \cdots \otimes |\xi_{m_d}^{\{d\}}\rangle. \tag{63}$$

$C \in \mathbb{C}^{r_1 \times \ldots \times r_d}$ is a reduced core tensor, that is hopefully much smaller than the original coefficient tensor, due to the optimal choice of basis.

For exact recovery, obtaining the basis vectors in the discrete setting is relatively straightforward. It can be achieved by applying a singular value decomposition (SVD) in every mode

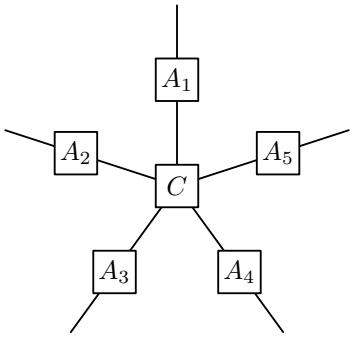


Figure 3: A Tucker tensor of order 5.

- thus called Higher Order SVD (HOSVD) - of the tensor: For the i-th mode, we compute the SVD of the *i-mode matricization*

$$[U]^{\alpha_i}_{\alpha_1,\ldots,\cancel{\alpha_i},\ldots,\alpha_d} \in \mathbb{C}^{(q_1\cdots\cancel{q_i}\cdots q_d)\times q_i} \tag{64}$$

and obtain the basis vectors $|\xi_1^{\{i\}}\rangle, \ldots, |\xi_{r_i}^{\{i\}}\rangle$, which span the optimal subspace of Λ_i[82,121,141].

In many applications, we want to approximate the tensor with lower rank $\tilde{\mathbf{r}} \leq \mathbf{r}$. In the matrix case $d = 2$, this can be done by truncating the above SVD and omitting the basis vectors that belong to the smallest $r - \tilde{r}$ singular values. The discovery that this already yields the optimal result is mostly accredited to Eckard and Young in mathematics, while most physics articles recognize the fact that it had been proven by Schmidt long before for the more complicated case of integral operators[209].

Unfortunately, this result cannot be generalized to tensors with $d > 2$. It has been shown in[97] that even finding the best rank one, i.e. $\mathbf{r} = (1, 1, \ldots, 1)$, can be NP-hard if $d > 2$. Nevertheless, truncating the HOSVD in every mode only yields a *quasi-optimal approximation* with respect to the l_2-norm[121]. However, in many cases, this is satisfactory.

The Tucker format is a subspace decomposition as the tensor is expressed in the basis of a subspace of the tensor space. At the same time, it yields a tensor tree, i.e. its representation fulfills criterion 3.1, 3.2 and 3.3. $C \in \mathbb{C}^{r_1\times\ldots\times r_d}$ is the only virtual component and

$$A_i(\alpha_i, m_i) = \langle \phi_{\alpha_i}^{\{i\}} | \xi_{m_i}^{\{i\}} \rangle \tag{65}$$

yields the d physical components, see Fig. 3.

The HOSVD gives us a constructive algorithm that computes the Tucker decomposition, i.e. a representation of the form (63) with minimal rank $\mathbf{r}$, in polynomial time. Additionally, the set of tensors with Tucker rank at most $\mathbf{r}$ is known to be Zariski-closed[120]. Therefore, it is closed and we overcome the border rank problem. In terms of storage complexity however, this format is far from being optimal. It now scales exponentially in r_i, i.e for $r := \max\{r_i\}$ the scaling is in $\mathcal{O}(dqr + r^d)$. Especially for small q_i, where we do not have $r_i \ll q_i$, we cannot hope for much reduction of complexity. In particular, for $q_i = 2$ we do not gain any nontrivial progress.

3.4 Matricization and tensor multiplication

To a certain extent, these representation allow to apply matrix analysis techniques to tensors. We therefore generalize the aforementioned matricization. Let $t \subseteq \{1, \dots, d\}$ be a collection of physical dimensions and $t^c := \{1, \dots, d\} \setminus t$ its complement. Then

$$[U]_{\alpha_t}^{\alpha_{t^c}} \in \mathbb{C}^{q_t} \otimes \mathbb{C}^{q_{t^c}} \tag{66}$$

is the matricization with $q_t = \{q_i \in \{q_1, \dots, q_d\} : i \in t\}$ as row dimensions and $q_{t^c} = \{q_i \in \{q_1, \dots, q_d\} \setminus q_t\}$ as column dimensions. A special case is the i-th matricization

$$[U]_{\alpha_1,\dots,\alpha_i}^{\alpha_{i+1},\dots,\alpha_d} \in \mathbb{C}^{q_1 \cdots q_i} \otimes \mathbb{C}^{q_{i+1} \cdots q_d} \tag{67}$$

utilized further down that is casting the first i-variables into the row index, and the remaining $d - i$ in the column index.

This Einstein-like notation allows us to introduce a *tensor multiplication*. Let $U \in \mathbb{C}^{q_{1,1} \times \dots \times q_{1,d_1}}$ and $V \in \mathbb{C}^{q_{2,1} \times \dots \times q_{2,d_2}}$. Then if for to matricizations $t_1 \in \{1, \dots, d_1\}, t_2 \in \{1, \dots, d_2\}$ it holds $q_{1,t_1^c} = q_{2,t_2} =: q_{t_1,t_2}$ we get

$$[U]_{\alpha_{1,t_1}}^{\alpha_{t_1,t_2}} [V]_{\alpha_{t_1,t_2}}^{\alpha_{2,t_2^c}} = \sum_{\alpha_{t_1,t_2}} U(\alpha_{1,t_1}, \alpha_{t_1,t_2}) V(\alpha_{t_1,t_2}, \alpha_{2,t_2^c}). \tag{68}$$

This is exactly the matrix multiplication of the matricizations and it is the contraction over the indeces α_{t_1,t_2}. In the case where no dual space is involved, i.e. no contraction is performed, we obtain the tensor product

$$[U]_{\alpha_{1,1},\dots,\alpha_{1,d_1}} [V]_{\alpha_{2,1},\dots,\alpha_{2,d_2}} = U \otimes V. \tag{69}$$

Note that in the complex case described here, the matricization should only be seen as the reordering and grouping of indeces, instead of introducing a concept of duality as done in some literature[119]. This is due to the fact that it is impossible to take the complex conjugate only in a few indeces of U, which would be required for this concept[227]. Thus, the reader should note that switching the ordering of the indeces gives only the transpose and not the hermitian of the original matricization:

$$[U]_{\alpha_t}^{\alpha_{t^c}} = \left([U]_{\alpha_{t^c}}^{\alpha_t}\right)^{\mathsf{T}} = \overline{\left([U]_{\alpha_{t^c}}^{\alpha_t}\right)^{\dagger}}. \tag{70}$$

Finally, we want to simplify the notation for the unambiguous case where we multiply over *all* common indeces. This will be denoted with a circle, since it can be seen as a composition of two linear operators:

$$U \circ V := [U]_{\alpha_{1,t_1}}^{\alpha_{t_1,t_2}} [V]_{\alpha_{t_1,t_2}}^{\alpha_{2,t_2^c}}, \; q_{1,i} \neq q_{2,j} \; \forall \; i \in t_1, j \in t_2^c. \tag{71}$$

3.5 Matrix product states or the tensor train format

Another example of a tensor network is the *Tensor Train (TT)* decomposition of a tensor. The tensor U is given element-wise as

$$U(\alpha_1, \dots, \alpha_d) = \sum_{m_1=1}^{r_1} \dots \sum_{m_{d-1}=1}^{r_{d-1}} A_1(\alpha_1, m_1) A_2(m_1, \alpha_2, m_2) \cdots A_d(m_{d-1}, \alpha_d). \tag{72}$$

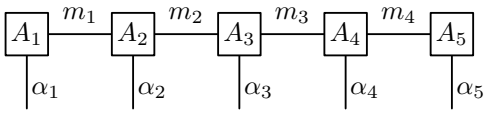


Figure 4: A tensor of order 5 in TT representation.

We get d component tensors of order 2 or 3. Their graph has the structure of a chain or train, hence the name. Figure (4) gives an example of a TT tensor.

The TT format maintains the positive characteristics of the Tucker format and overcomes most of the disadvantages of the canonical format. However, the complexity now scales only quadratically in the ranks, or with $\mathcal{O}(qdr^2)$, for $r = \max\{r_i\}$. While the Tensor Train decomposition is not the only format that has this advantage, it is one of the most widely used ones and it will also be the standard format in this paper.

This format has been introduced to the mathematical realm by Oseledets et al.[182]. While it was developed independently, it can be seen as a special case of the *Hierarchical Tucker (HT)* decomposition. However, we will restrict ourselves to the TT format and deal with the HT format only briefly further down. As stated above, nearly everything of the following can be generalized to a general tensor tree format without effort, but notation becomes more complex.

In physics the Tensor Train decomposition has been known as Matrix Product States (MPS) since the late nineties and many results can be taken directly from there. The name Matrix Product States is justified if we fix the physical indices. This yields a chain of matrix products:

$$U(\alpha_1, \ldots, \alpha_d) = \mathbf{A}_1(\alpha_1)\mathbf{A}_2(\alpha_2)\cdots\mathbf{A}_{d-1}(\alpha_{d-1})\mathbf{A}_d(\alpha_d) \tag{73}$$

with $[\mathbf{A}_i(\alpha_i)]_{m_{i-1},m_i} := A_i(m_{i-1}, \alpha_i, m_i) \in \mathbb{C}^{r_{i-1}\times r_i}$.

Let it be noted that an important modification of the Tensor Train format follows if we introduce a contraction of rank greater than 1 between the first and last component, also called *periodic boundary conditions*,

$$U(\alpha_1, \ldots, \alpha_d) = \sum_{m_1=1}^{r_1} \cdots \sum_{m_d=1}^{r_d} A_1(m_d, \alpha_1, m_1)A_2(m_1, \alpha_2, m_2)\cdots A_d(m_{d-1}, \alpha_d, m_d). \tag{74}$$

These *uniform Matrix Product States (uMPS)* are especially significant in physics. Verstraete et al. deal with uMPS that are also translation invariant, i.e. all components are equal $A_1 = \ldots = A_d$[185]. The graph of this decomposition is circular and therefore does not suffice criterion 3.3. As mentioned above, this poses a number of problems[119] that are - in a nutshell - similar to those of the canonical format. For this reason, we will only deal with regular, non-circular Matrix Product States from now on.

The TT format can be considered as a multi-layered subspace representation. This is achieved in a hierarchical way[82]. In Λ_1 we consider the subspace Ξ_1 given by the basis set $\{|\xi^{\{1\}}_{m_1}\rangle : m_1 = 1, \ldots, r_1\}$, where

$$|\xi^{\{1\}}_{m_1}\rangle := \sum_{\alpha_1=1}^{q_1} A_1(\alpha_1, m_1)\ |\phi^{\{1\}}_{\alpha_1}\rangle. \tag{75}$$

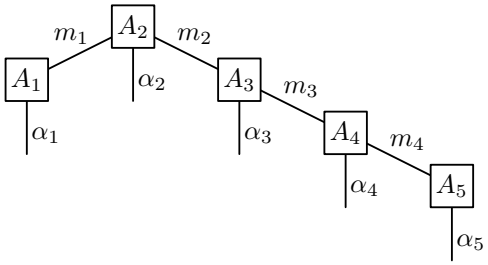


Figure 5: Hierarchical picture of a Tensor Train with A_2 as the root.

We proceed with a subspace of the partial tensor product space $\Xi_{\{1,2\}} \subset \Lambda_1 \otimes \Lambda_2$ of dimension $r_{\{1,2\}} \leq q_1 q_2$. Indeed $\Xi_{\{1,2\}}$ is defined through a new basis set $\{|\xi^{\{1,2\}}_{m_{\{1,2\}}}\rangle : 1, \dots, r_{\{1,2\}}\}$ where the new basis vectors are given in the form

$$|\xi^{\{1,2\}}_{m_{\{1,2\}}}\rangle = \sum_{\alpha_1=1}^{q_1} \sum_{\alpha_2=1}^{q_2} U_{\{1,2\}}(\alpha_1, \alpha_2, m_{\{1,2\}}) \; |\phi^{\{1\}}_{\alpha_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle. \tag{76}$$

We observe that $\Xi_{\{1,2\}} \subset \Xi_1 \otimes \Lambda_2$ with

$$|\xi^{\{1,2\}}_{m_{\{1,2\}}}\rangle = \sum_{m_1=1}^{r_1} \sum_{\alpha_2=1}^{q_2} A_2(m_1, \alpha_2, m_{\{1,2\}}) \; |\xi^{\{1\}}_{m_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle \tag{77}$$

and thus

$$U_{\{1,2\}}(\alpha_1, \alpha_2, m_{\{1,2\}}) = \sum_{m_1=1}^{r_1} A_1(\alpha_1, m_1) A_2(m_1, \alpha_2, m_{\{1,2\}}). \tag{78}$$

For this reason, when dealing with TT tensors, we simplify the notation and often set $\{1,2\} \simeq 2$, and in general $\{1,2,\dots,i\} \simeq i$.

The tensor is recursively defined by the component tensors A_i,

$$\begin{aligned} |\xi^{\{3\}}_{m_3}\rangle &= \sum_{m_2,\alpha_3} A_3(m_2, \alpha_3, m_3) \; |\xi^{\{2\}}_{m_2}\rangle \otimes |\phi^{\{3\}}_{\alpha_3}\rangle \\ &= \sum_{m_1,m_2,\alpha_2,\alpha_3} A_2(m_1, \alpha_2, m_2) A_3(m_2, \alpha_3, m_3) \; |\xi^{\{1\}}_{m_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle \otimes |\phi^{\{3\}}_{\alpha_3}\rangle \\ &= \sum_{m_1,m_2,\alpha_1,\alpha_2,\alpha_3} A_1(\alpha_1, m_1) A_2(m_1, \alpha_2, m_2) A_3(m_2, \alpha_3, m_3) \; |\phi^{\{1\}}_{\alpha_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle \otimes |\phi^{\{3\}}_{\alpha_3}\rangle, \end{aligned} \tag{79}$$

and so forth, by taking $\Xi_{\{1,\dots,i+1\}} \subset \Xi_{\{1,\dots,i\}} \otimes \Lambda_{i+1}$.

We may also proceed differently, e.g. $\Xi_{\{1,2,3,4,\dots\}} \subset \Xi_{\{1,2\}} \otimes \Xi_{\{3,4\}} \otimes \dots$. Especially, it can be advantageous to start from the right hand side, i.e. taking $\Lambda_i \otimes \Xi_{\{i+1,\dots,d\}}$ etc., obtaining basis vectors

$$|\zeta^{\{i\}}_{m_{i-1}}\rangle \in \Xi_{\{i,\dots d\}}. \tag{80}$$

Let us fix some $i \in \{1, \dots, d\}$ and call it the *root*. This gives a hierarchical picture (see Fig. 5).

We consider the spaces $\mathcal{L}_i := \Xi_{\{1,\dots,i-1\}}$ and $\mathcal{R}_j := \Xi_{\{i+1,\dots,d\}}$. Their dimensions are given by

$$\dim \mathcal{L}_i = r_{i-1}, \dim \mathcal{R}_i = r_i \tag{81}$$

and hence, the full tensor $|\Psi\rangle$ is contained in the $r_{i-1}q_ir_i$-dimensional subspace[77,84]

$$|\Psi\rangle \in \mathcal{L}_i \otimes \Lambda_i \otimes \mathcal{R}_i \cong \mathbb{C}^{r_{i-1}\times q_i \times r_i} \tag{82}$$

$$|\Psi\rangle = \sum_{m_{i-1},m_i,\alpha_i} A_i(m_{i-1},\alpha_i,m_i)\ |\xi^{\{i-1\}}_{m_{i-1}}\rangle \otimes |\phi^{\{i\}}_{\alpha_i}\rangle \otimes |\zeta^{\{i+1\}}_{m_i}\rangle \tag{83}$$

A canonical but not necessary choice is that the basis vectors $|\xi^{\{i-1\}}_1\rangle, \dots, |\xi^{\{i-1\}}_{r_{i-1}}\rangle$ and $|\zeta^{\{i+1\}}_1\rangle, \dots, |\zeta^{\{i+1\}}_{r_i}\rangle$ are orthogonal and normalized.

We will see in the following that this hierarchical or multi-layered subspace approximation constitutes the mechanism behind the *renormalization group* formalism in the one-site DMRG (density matrix renormalization group).

An obvious observation[100] following from the above will be that the minimal dimension r_i is the rank of the i-th matricization (67):

Theorem 3.4 (Separation Theorem). *For any tensor $U \in \mathbb{C}^{q_1\times\cdots\times q_d}$, there exists a minimal TT (MPS) representation, thus called TT decomposition TT(U), such that for any $i = 1,\dots,d-1$ the dimensions r_i of the contractions $m_i = 1,\dots,r_i$ are minimal and given by*

$$r_i = \operatorname{rank}([U]^{\alpha_{i+1},\dots,\alpha_d}_{\alpha_1,\dots,\alpha_i}). \tag{84}$$

We can change the hierarchy, e.g. by choosing the next component A_{i+1} as the root. In most applications, it will then become necessary to shift the orthogonalization such that $\{|\xi^{\{i\}}_{m_i}\rangle : m_i = 1,\dots,r_i\}$ and $\{|\zeta^{\{i+2\}}_{m_{i+1}}\rangle : m_{i+1} = 1,\dots,r_{i+1}\}$ are orthonormal. This can be done by applying the singular value decomposition to the matricization of the i-th component

$$[A_i]^{m_{i+1}}_{m_{i-1},\alpha_i} = [\tilde{A}_i]^{m_{i+1}}_{m_{i-1},\alpha_i}\mathbf{\Sigma}_i\mathbf{Y}_i^\dagger \tag{85}$$

and shifting $\mathbf{\Sigma}_i, \mathbf{Y}_i^\dagger \in \mathbb{C}^{r_i\times r_i}$ to the next component

$$[\tilde{A}_{i+1}]^{\alpha_{i+1},m_{i+2}}_{m_i} = \mathbf{\Sigma}_i\mathbf{Y}_i^\dagger[A_{i+1}]^{\alpha_{i+1},m_{i+2}}_{m_i}. \tag{86}$$

For $|\Psi\rangle$ we obtain

$$\begin{aligned}
|\Psi\rangle &= \sum_{m_{i-1},m_i,\alpha_i} A_i(m_{i-1},\alpha_i,m_i)\ |\xi^{\{i-1\}}_{m_{i-1}}\rangle \otimes |\phi^{\{i\}}_{\alpha_i}\rangle \otimes |\zeta^{\{i+1\}}_{m_i}\rangle \\
&= \sum_{\substack{m_{i-1},m_i,m_{i+1}\\ \alpha_i,\alpha_{i+1}}} A_i(m_{i-1},\alpha_i,m_i)A_{i+1}(m_i,\alpha_{i+1},m_{i+1})\ |f\xi^{\{i-1\}}_{m_{i-1}}\rangle \otimes |\phi^{\{i\}}_{\alpha_i}\rangle \otimes |\phi^{\{i+1\}}_{\alpha_{i+1}}\rangle \otimes |\zeta^{\{i+2\}}_{m_{i+1}}\rangle \\
&= \sum_{\substack{m_{i-1},m_i,m_{i+1}\\ \alpha_i,\alpha_{i+1}}} \tilde{A}_i(m_{i-1},\alpha_i,m_i)\tilde{A}_{i+1}(m_i,\alpha_{i+1},m_{i+1})\ |\xi^{\{i-1\}}_{m_{i-1}}\rangle \otimes |\phi^{\{i\}}_{\alpha_i}\rangle \otimes |\phi^{\{i+1\}}_{\alpha_{i+1}}\rangle \otimes |\zeta^{\{i+2\}}_{m_{i+1}}\rangle \\
&= \sum_{m_i,m_{i+1},\alpha_{i+1}} \tilde{A}_{i+1}(m_i,\alpha_{i+1},m_{i+1})\ |\xi^{\{i\}}_{m_i}\rangle \otimes |\phi^{\{i+1\}}_{\alpha_{i+1}}\rangle \otimes |\zeta^{\{i+2\}}_{m_{i+1}}\rangle.
\end{aligned} \tag{87}$$

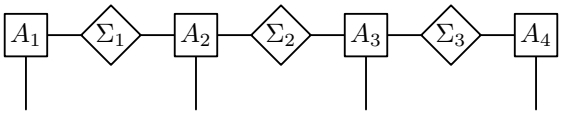


Figure 6: A TT tensor of order 4 in standard representation.

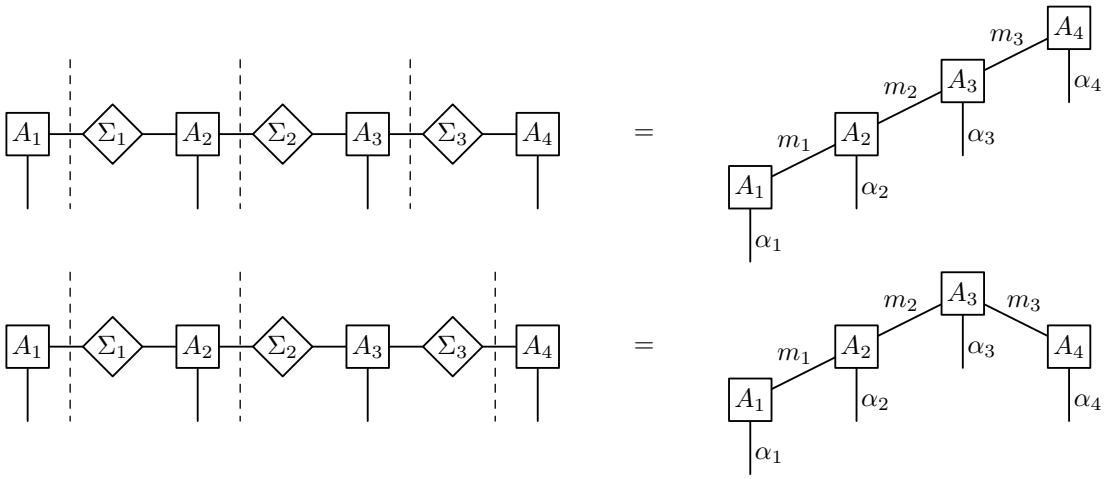


Figure 7: A shift of orthogonality in the standard representation.

Alternatively one may use QR factorization for the orthogonalization, but it often advantageous to keep the small diagonal matrix $\mathbf{\Sigma}_i \in \mathbb{C}^{r_i \times r_i}$ containing the singular values in between to adjacent component tensors. In fact this provides a *standard representation* or *HSVD representation* of U, see Fig. 6

$$U = A_1 \circ \mathbf{\Sigma}_1 \circ A_2 \circ \mathbf{\Sigma}_2 \circ \cdots \circ \mathbf{\Sigma}_{d-1} \circ A_d. \tag{88}$$

This representation has been developed independently by different authors[76,183,243]. In physics, it is accredited to Vidal and is hence also known as the *Vidal representation*. Very beneficial is the criterion that

$$A_1^\dagger A_1 = \mathbf{I}_{r_1}, \;\; A_d A_d^\dagger = \mathbf{I}_{r_{d-1}} \tag{89}$$

and for all $1 < i < d$

$$[A_i \circ \mathbf{\Sigma}_i]_{m_{i-1}}^{\alpha_i, m_i} \big([A_i \circ \mathbf{\Sigma}_i]_{m_{i-1}}^{\alpha_i, m_i}\big)^\dagger = \mathbf{I}_{r_{i-1}}, \tag{90}$$

$$\big([\mathbf{\Sigma}_{i-1} \circ A_i]_{m_{i-1}, \alpha_i}^{m_i}\big)^\dagger [\mathbf{\Sigma}_{i-1} \circ A_i]_{m_{i-1}, \alpha_i}^{m_i} = \mathbf{I}_{r_i}. \tag{91}$$

This means, we can shift the root, and thus the orthogonality, by simply shifting the *density matrices* $\mathbf{\Sigma}_i$, see Fig. 7.

This representation can be computed by applying a sequence of singular value decomposition and storing the singular values. The procedure is called *Hierarchical SVD (HSVD)*. It recovers the tensor exactly. However, as mentioned for the Tucker format and the HOSVD, the HSVD can be used for approximation by thresholding the singular values. For density

matrices $\mathbf{\Sigma} = \operatorname{diag}(\sigma_1, \dots, \sigma_r)$ we define two thresholding operators

$$H_{\tilde{r}}(\mathbf{\Sigma}) = \operatorname*{diag}_{1 \le i \le r}(\sigma_i) \ , \ \tilde{r} \le r, \tag{92}$$

$$H_{\epsilon}(\mathbf{\Sigma}) = \operatorname*{diag}_{\sigma_i \ge \epsilon}(\sigma_i) \ , \ \epsilon > 0 \tag{93}$$

and for TT tensors

$$H_{\tilde{\mathbf{r}}}(U) = A_1 \circ H_{\tilde{r}_1}(\mathbf{\Sigma}_1) \circ A_2 \circ H_{\tilde{r}_2}(\mathbf{\Sigma}_2) \circ \cdots \circ H_{\tilde{r}_{d-1}}(\mathbf{\Sigma}_{d-1}) \circ A_d, \tag{94}$$

$$H_{\epsilon}(U) = A_1 \circ H_{\epsilon_1}(\mathbf{\Sigma}_1) \circ A_2 \circ H_{\epsilon_2}(\mathbf{\Sigma}_2) \circ \cdots \circ H_{\epsilon_{d-1}}(\mathbf{\Sigma}_{d-1}) \circ A_d. \tag{95}$$

Again, this will not yield the best approximation of the tensor, as it does in the matrix case. As with Tucker tensors, we maintain a so called *quasi optimality*:

Theorem 3.5 (Quasi Optimality). [76,82,83,183] *The truncation of the HSVD can be estimated by*

$$\|U - H_{\tilde{\mathbf{r}}}(U)\| \le \sqrt{d-1} \inf_{V \in \mathcal{M}_{\le \tilde{\mathbf{r}}}} \|U - V\|, \tag{96}$$

where $\mathcal{M}_{\le \tilde{\mathbf{r}}}$ is the space of all tensors with TT rank not exceeding $\tilde{\mathbf{r}}$.

As most other results, the separation theorem and the quasi optimality can be readily generalized to all tree tensor networks. It is also possible to formulate a standard representation for other trees. In contrast to the parametrization (58), the subspace representation provides further essential information about minimal representability and approximability. It justifies the use of the notion of *entanglement* for the tensor U or an appropriate low rank approximation of it. Entanglement here means the quantum correlation between the subsystem consisting of the first i orbitals and the subsystem consisting of the remaining orbitals, and it can be characterized by, e.g., the quantum Hartley entropy $\log r_i$, see in section 4.2.1). Without further conditions, these quantities are not well defined for tensor representations that do not have a tree structure. However, Verstraete developed an *injectivity condition* that aims at overcoming that problem for uniform MPS with periodic boundary conditions[87].

3.6 Dimension trees and the hierarchical tensor decomposition

We briefly discuss the *Hierarchical Tucker* (HT) representation that has been introduced by Hackbusch and Kuhn[81] in 2009 and has since received a lot of attention. This is also due to the fact that it is a reasonable generalization of the TT format.

The HT representation is defined by a *dimension tree*, usually a binary tree, where the leafs $U_{\{1\}}, \dots, U_{\{d\}} = A_1, \dots, A_d$ constitute the physical components and the inner vertices U_t are virtual. Hackbusch gives the following comprehensive notation in[82]: The vertices of the tree tensor network $TTNS(U) = (V, E, H)$ are labeled

i) $t_r = \{1, \dots, d\}$ for the root,

ii) $t \in L := \big\{\{1\}, \dots, \{d\}\big\}$ for the leafs and

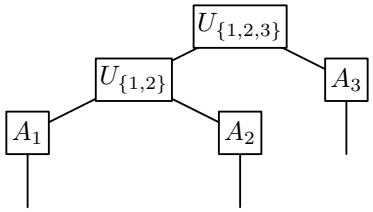


Figure 8: A tensor of order 3 in HT format.

iii) $t \in V \setminus L$ for inner vertices, which have sons $t_1, \dots, t_p$ that are an ordered partition of t, i.e.

$$\bigcup_i^p t_i = t \text{ and } \mu < \nu \; \forall \; \mu \in t_i, \; \nu \in t_j, \; i < j.$$

For an inner vertex $t \subset V \setminus L$, with sons $t_1, \dots, t_p$ (usually $p = 2$), there is a subspace Ξ_t defined by its basis set $\{|\xi_{m_t}^{\{t\}}\rangle : m_t = 1, \dots, r_t\}$[68] given by

$$|\xi_{m_t}^{\{t\}}\rangle = \sum_{m_1=1}^{r_{t_1}} \cdots \sum_{m_p=1}^{r_{t_p}} U_t(m_1, \dots, m_p, m_t) \; |\xi_{m_1}^{\{t_1\}}\rangle \otimes \cdots \otimes |\xi_{m_p}^{\{t_p\}}\rangle. \tag{97}$$

The root $t_r = \{1, \dots, d\}$, with sons $t_1, \dots, t_p$, is to reconstruct the tensor

$$U = \sum_{m_1=1}^{r_{t_1}} \cdots \sum_{m_p=1}^{r_{t_p}} U_{t_r}(m_1, \dots, m_p) \; |\xi_{m_1}^{\{t_1\}}\rangle \otimes \cdots \otimes |\xi_{m_p}^{\{t_p\}}\rangle. \tag{98}$$

Therefore the tensor U is defined completely by the component tensors U_t, using the above representations recursively, see Fig. 8. There are at most $\mathcal{O}(d)$ vertices and consequently the complexity is $\mathcal{O}(qdr + dr^{p+1})$. For $p = 2$ we obtain $\mathcal{O}(qdr + dr^3)$[82,84].

As with the Tucker and the TT format, obtaining the HT format can be done by applying the singular value decomposition successively, in a hierarchical fashion. Again, we maintain a well-defined rank through a separation theorem, a quasi optimality of a truncated HSVD and so on.

In fact, the Tensor Train decomposition can be seen as a special case of the Hierarchical Tucker decomposition, where we use an unbalanced tree and omit the optimal subspaces in the leafs. However, in some cases, the binary tree structure can be advantageous[244]. Additionally, the leafs $A_1, \dots, A_d$ form exactly the optimal subspaces already observed in the Tucker decomposition. We refer to the literature cited above.

Arguably, this could make the HT format superior to the TT format. However, the notation becomes very messy and all notable theoretical results are valid for any tree tensor network. Hence, we refrain from dealing with the Hierarchical format and proceed with the Tensor Train format, keeping the similarities in mind.

3.7 Fixed rank manifolds and varieties

For many applications, we consider the set of tensors of fixed TT rank

$$\mathcal{M}_{\mathbf{r}} := \{U \in \mathbb{C}^{n_1 \times \dots \times n_d} : \mathbf{r}_{TT} = \mathbf{r}\}.$$

This set is no longer a linear space nor is it convex. In order to parametrize this space, we introduce the component set $\mathcal{C} = \{\underline{U} = (A_1, \dots, A_d) : A_i \in \mathbb{C}_*^{r_{i-1} \times q_i \times r_i}\}$ and the map

$$\tau : \mathcal{C} \to \mathcal{M}_{\mathbf{r}} \subseteq \mathbb{C}^{q_1 \times \dots \times q_d}, \tag{99}$$

$$(A_1, \dots, A_d) \mapsto \tau(A_1, \dots, A_d) := U. \tag{100}$$

For each i, $\mathbb{C}_*^{r_{i-1} \times q_i \times r_i}$ is the space of all elements with full multilinear rank:

$$\mathbb{C}_*^{r_{i-1} \times q_i \times r_i} := \{A_i \in \mathbb{C}^{r_{i-1} \times q_i \times r_i} : \operatorname{rank}([A_i]_{m_{i-1}}^{\alpha_i, m_i}) = r_{i-1},\ \operatorname{rank}([A_i]_{m_{i-1}, \alpha_i}^{m_i}) = r_i\} \tag{101}$$

Let it be noted that this space is a smooth manifold[7].

The map τ is clearly surjective onto $\mathcal{M}_{\mathbf{r}}$, but it is not injective: For any non-singular matrix $\mathbf{X} \in GL(r_i) \subseteq \mathbb{C}^{r_i \times r_i}$ we have

$$A_1 \circ \dots \circ A_d = A_1 \circ \dots \circ A_i \circ \mathbf{X} \circ \mathbf{X}^{-1} \circ A_{i+1} \circ \dots \circ A_d. \tag{102}$$

Any parametrization of the form (58) shares this kind of non-uniqueness. But for tree tensor networks, this problem can be overcome: Let $\mathcal{G}$ be the Lie group

$$\mathcal{G} = \{g = (\mathbf{X}_1, \dots, \mathbf{X}_{d-1}) : \mathbf{X}_i \in \mathrm{GL}(r_i)\} \cong \bigotimes_{i=1}^{d-1} \mathrm{GL}(r_i). \tag{103}$$

We define the group action of $g \in \mathcal{G}$ on the components $\underline{U}$ as

$$g \cdot \underline{U} := (A_1 \circ \mathbf{X}_1, \mathbf{X}_1^{-1} \circ A_2 \circ \mathbf{X}_2, \dots, \mathbf{X}_{d-1}^{-1} \circ A_d). \tag{104}$$

This action is smooth and it acts freely and properly on $\mathcal{C}$, see[232]. The orbits are the equivalence classes, given by

$$[\underline{U}] = \mathcal{G} \cdot \underline{U} = \{g \cdot \underline{U} : g \in \mathcal{G}\}. \tag{105}$$

Thus, we obtain the quotient space

$$\mathcal{C}/\mathcal{G} = \{[\underline{U}] = \mathcal{G} \cdot \underline{U} : \underline{U} \in \mathcal{C}\} \tag{106}$$

with the quotient map

$$\pi : \mathcal{C} \to \mathcal{C}/\mathcal{G},\ \underline{U} \mapsto [\underline{U}]. \tag{107}$$

This yields a bijection

$$\hat{\tau} : \mathcal{C}/\mathcal{G} \to \mathcal{M}_{\mathbf{r}}, \tag{108}$$

where $\tau = \hat{\tau} \circ \pi$. As a result, we get that $\mathcal{M}_{\mathbf{r}}$ is a smooth quotient manifold[122].

This manifold can be globally embedded into the tensor space $\mathcal{M}_{\mathbf{r}} \subset \mathbb{C}^{q_1 \times \dots \times q_d}$ and we call it the *TT manifold*[13,85,88,100,232]. Thus, it is possible to define the tangent space $\mathrm{T}_U \mathcal{M}_{\mathbf{r}}$, which is a linear subset of $\mathbb{C}^{q_1 \times \dots \times q_d}$. It is isomorphic to the *horizontal space*

$$\mathrm{H}_{\underline{U}} \mathcal{C} = \{(W_1, \dots, W_d) \in \mathcal{C} : \left([W_i]_{k_{i-1}, x_i}^{k_i}\right)^\dagger [A_i]_{k_{i-1}, x_i}^{k_i} = \mathbf{0}\ \forall\ i = 1, \dots, d-1\} \tag{109}$$

via

$$\big(D\tau(\underline{U})\big)(W_1,\dots,W_d) = \sum_{i=1}^{d} A_1 \circ \cdots \circ W_i \circ \cdots \circ A_d. \tag{110}$$

We remark that different definitions of the horizontal space are possible and that the choice of the gauge conditions above is not unique. It also depends on the choice of the root. In the above case, the root is set to be the last component A_d. The only requirement for a horizontal space is that it forms the tangent space of $\mathcal{C}$ via the direct sum

$$\mathrm{T}_{\underline{U}}\mathcal{C} = \mathrm{V}_{\underline{U}}\mathcal{C} \oplus \mathrm{H}_{\underline{U}}\mathcal{C}, \tag{111}$$

where $\mathrm{V}_{\underline{U}}\mathcal{C}$ is the *vertical space* tangential to the orbits.

The manifold $\mathcal{M}_{\mathbf{r}}$ is an open set. However, in finite dimensions, its closure is given by

$$\overline{\mathcal{M}_{\mathbf{r}}} = \mathcal{M}_{\leq \mathbf{r}}. \tag{112}$$

This is based on the observation that the matrix rank is an upper semi-continuous function[68,82]. The singular points are exactly those where the actual rank is not maximal.

As mentioned above, the set $\mathcal{M}_{\leq \mathbf{r}}$ is Zariski-closed and thus forms an *algebraic variety*, i.e. it is the set of common zeros of polynomials. This is easy to see: Indeed, we know from the separation theorem, that $\mathcal{M}_{\leq \mathbf{r}}$ is the intersection of all tensors where the corresponding matricizations $[U]^{\alpha_{i+1},\dots,\alpha_d}_{\alpha_1,\dots,\alpha_i}$ have at most rank r_i. The sets of matrices with rank at most r_i are known to be algebraic varieties[210], each some zero-set of polynomials[119]. Then, trivially, the intersection is the zero-set of the union of all such polynomials. Again, this property generalizes to all tensor trees.

3.8 Dirac-Frenkel variational principle or dynamical low rank approximation

Solving problems in the large tensor space is often too expensive due to the curse of dimensionality. We therefore restrict ourselves to tensors of fixed rank, i.e. to the space $\mathcal{M}_{\leq \mathbf{r}}$. In general, the appropriate ranks are unknown. Thus, we start with an initial guess and increase the ranks when necessary. There are some greedy techniques available that serve this purpose[68].

For the approximation with fixed rank, we consider the smooth manifold $\mathcal{M}_{\mathbf{r}}$, as this facilitates the theoretical framework. Let

$$J(U) \to \min,\ \ J \in C^1(\mathbb{C}^{q_1,\dots,q_d}, \mathbb{R}) \tag{113}$$

be a minimization problem on the tensor space, for example the minimization of the energy functional (45a) in quantum chemistry.

For the restriction of J to $\mathcal{M}_{\mathbf{r}}$ we obtain the necessary condition

$$U = \operatorname*{argmin}_{V \in \mathcal{M}_{\mathbf{r}}} J(V) \Rightarrow \langle \nabla J(V), \delta U \rangle = 0 \ \forall\ \delta U \in \mathrm{T}_U \mathcal{M}_{\mathbf{r}}, \tag{114}$$

i.e. if U minimizes J on $\mathcal{M}_{\mathbf{r}}$, then the gradient of J must be orthogonal to the tangent space at U. Equivalently, if we denote the orthogonal projection onto $\mathrm{T}_U\mathcal{M}_{\mathbf{r}}$ with P_{T_U}, we get

$$P_{\mathrm{T}_U} \nabla J(U) = 0. \tag{115}$$

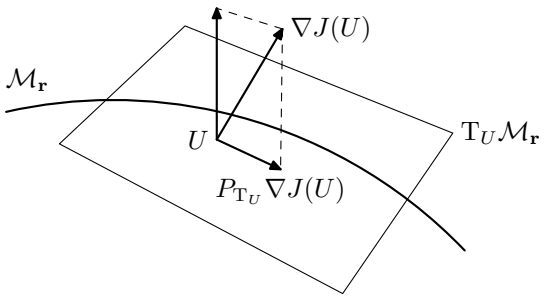


Figure 9: An illustration of the gradient flow on manifold $\mathcal{M}_{\mathbf{r}}$.

This variational approach can be generalized to the dynamical problem

$$\frac{d}{dt}U = f(U), \tag{116}$$

$$U(0) = U_0 \in \mathcal{M}_{\mathbf{r}}. \tag{117}$$

This is a differential equation on $M_{\mathbf{r}}$ if and only if

$$f(U) \in \mathrm{T}_U\mathcal{M}_{\mathbf{r}},\ \forall U \in \mathcal{M}_{\mathbf{r}}. \tag{118}$$

Thus (116) can be solved approximately by projecting $f(U)$ on the tangent space $\mathrm{T}_U\mathcal{M}_{\mathbf{r}}$,

$$F(U) := P_{\mathrm{T}_U} f(U) \tag{119}$$

and solving the projected differential equation

$$\frac{d}{dt}U = F(U). \tag{120}$$

In accordance with the above, we obtain

$$\frac{d}{dt}U - F(U) = 0 \Leftrightarrow \langle \frac{d}{dt}U - f(U), \delta U\rangle = 0\ \forall\ \delta U \in \mathrm{T}_U\mathcal{M}_{\mathbf{r}}. \tag{121}$$

In the context of time-dependent quantum chemistry, this is well-known as the *Dirac-Frenkel variational principle*[20,49,91,140,141,156].

Replacing $f(U)$ with $-\nabla J(U)$ in (116) gives the gradient flow of J. Then (121) becomes

$$\langle \frac{d}{dt}U + \nabla J(U), \delta U\rangle = 0\ \forall\ \delta U \in \mathrm{T}_U\mathcal{M}_{\mathbf{r}} \tag{122}$$

and a solution can be computed with the aforementioned methods[13,54,84,87,140,142], see Fig. 9.

3.9 The alternating least squares algorithm

Consider the functional

$$j : \mathcal{C} \to \mathbb{R} \tag{123}$$

$$(A_1, \ldots, A_d) \mapsto j(A_1, \ldots, A_d) := J(\tau(A_1, \ldots, A_d)). \tag{124}$$

For $i \in \{1, \dots, d\}$ we fix $A_1, \dots, A_{i-1}$ and $A_{i+1}, \dots, A_d$ and solve the subproblem

$$A_i^+ := \operatorname*{argmin}_{V_i \in \mathbb{C}^{r_{i-1} \times q_i \times r_i}} j(A_1, \dots, V_i, \dots, A_d). \tag{125}$$

This is done in a successive manner and with alternating directions, which - for the best least squares fit $J(U) = \|U - B\|$ - justifies the name *Alternating Least Squares* (ALS) algorithm. The well-known Gauß-Seidel iteration is based on this strategy.

The TT format allows for a special formulation of this algorithm, sometimes dubbed the *Alternating Linear Scheme* to maintain the abbreviation. In this case, we can give a closed form for each subproblem and they can be solved using standard tools from linear algebra and numerical optimization.

In every step, one has to solve a small problem in order to achieve the minimum. Note that we allow $V_i \in \mathbb{C}^{r_{i-1} \times q_i \times r_i}$, i.e. the ranks can decrease in each step. This automatically restricts J to the variety $\mathcal{M}_{\leq \mathbf{r}}$ since the components can have full rank or less, but obviously not more than that.

The small subproblems will be of the same kind as the original problem, i.e. linear equations will be turned into small linear equations and eigenvalue problems give rise to relatively small (generalized) eigenvalue problems. In physics this supports the renormalization picture, where an original large systems is reduced to a small system with the same ground state energy and possibly further physical quantities.

As we have observed before, this simple approach should be realized with some care. Since the representation is redundant, we can generally not minimize over the full parameter space $\mathbb{C}^{r_{i-1} \times q_i \times r_i}$ but rather some non-linear quotient space and it becomes necessary to introduce gauge conditions like above. However, this can be avoided if we choose to minimize only the root of the tensor as there is no redundancy in this part. After the minimization, it would then be crucial to restructure the hierarchy of the tensor and consider the next component as the root. This can be done by shifting the orthogonality as explained in (87). The extension to general hierarchical trees is straightforward.

Conforming with the earlier notation (82), each subproblem becomes a problem over a small subset that constitutes a subspace

$$\mathcal{L}_i \otimes \Lambda_i \otimes \mathcal{R}_i \subseteq \Lambda^{(d)} = \bigotimes_{i=1}^{d} \Lambda_i. \tag{126}$$

We define the orthogonal projector onto this space

$$P_i : \Lambda^{(d)} \to \mathcal{L}_i \otimes \Lambda_i \otimes \mathcal{R}_i. \tag{127}$$

If we choose orthogonal bases for $|\xi_1^{\{i-1\}}\rangle, \dots, |\xi_{r_{i-1}}^{\{i-1\}}\rangle$ and $|\zeta_1^{\{i+1\}}\rangle, \dots, |\zeta_{r_i}^{\{i+1\}}\rangle$, we obtain

$$P_i \simeq E_i E_i^\dagger, \tag{128}$$

where

$$E_i : \mathbb{C}^{r_{i-1} \times q_i \times r_i} \to \mathbb{C}^{q_1 \times \dots \times q_d} \tag{129}$$

$$V_i \mapsto E_i V_i = A_1 \circ \dots \circ A_{i-1} \circ V_i \circ A_{i+1} \circ \dots \circ A_d \tag{130}$$

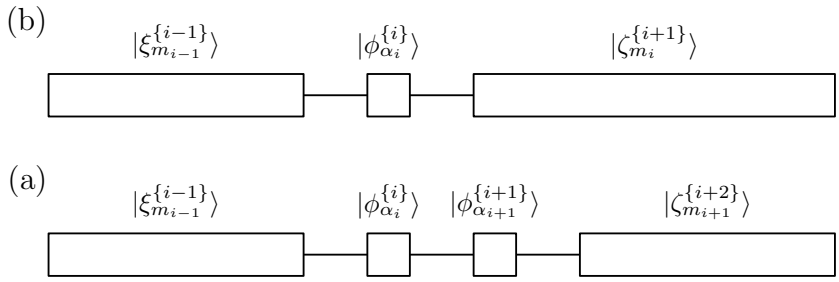


Figure 10: Reduced basis representation for (a) the ALS algorithm, and (b) the two-site DMRG.

is the *insertion operator* also use elsewhere[99]. This can easily be seen, as for $V \in \mathbb{C}^{q_1 \times \ldots \times q_d}$ it holds

$$P_i|\Psi\rangle = \sum_{m_{i-1},m_i,\alpha_i} \tilde{V}_i(m_{i-1},\alpha_i,m_i)\ |\xi^{\{i-1\}}_{m_{i-1}}\rangle \otimes |\phi^{\{i\}}_{\alpha_i}\rangle \otimes \zeta^{\{i+1\}}_{m_i}\rangle \tag{131}$$

$$= \sum_{\alpha_1,\ldots,\alpha_d} E_i \tilde{V}_i(\alpha_1,\ldots,\alpha_d)\ |\phi^{\{1\}}_{\alpha_1}\rangle \otimes \ldots \otimes |\phi^{\{d\}}_{\alpha_d}\rangle \tag{132}$$

and

$$E_i^\dagger V = \tilde{V}_i. \tag{133}$$

Note that E_i is a bijection onto its image, and since it is also orthogonal, its hermitian is well-defined as its inverse. See Fig. 10(a) for an illustration of the reduced basis.

To formulate the procedure explicitly, we consider a linear system, i.e. a functional

$$J(U) = \frac{1}{2}\langle \mathbf{X}U, U\rangle - \langle B, U\rangle, \tag{134}$$

where $\mathbf{X} \in \mathrm{Lin}(\mathbb{C}^{q_1 \times \ldots \times q_d}, \mathbb{C}^{q_1 \times \ldots \times q_d})$ is a linear operator. This operator can be stored and viewed in a canonical-like format, i.e. as a sum of rank-one tensor products

$$\mathbf{X} = \sum_k \mathbf{X}_1^k \otimes \cdots \otimes \mathbf{X}_d^k \tag{135}$$

or even in a TT-Matrix or Matrix Product Operator (MPO) format[182]. This is irrelevant for the purpose of notation, but it can be of computational interest.

Since we have equivalence $\Lambda^{(d)} \cong \mathbb{C}^{q_1 \times \ldots \times q_d}$, we also denote $\mathbf{X} \in \mathrm{Lin}(\Lambda^{(d)}, \Lambda^{(d)})$ without changing the notation. For the right side we denote $B \simeq |\boldsymbol{\Upsilon}\rangle \in \Lambda^{(d)}$. A single subproblem can then be expressed as

$$\begin{aligned} A_i^+ &= \operatorname*{argmin}_{V_i \in \mathbb{C}^{r_{i-1} \times q_i \times r_i}} j(A_1,\ldots,V_i,\ldots,A_d) \\ &= \operatorname*{argmin}_{V_i \in \mathbb{C}^{r_{i-1} \times q_i \times r_i}} \big(\frac{1}{2}\langle \mathbf{X}E_iV_i, E_iV_i\rangle - \langle B, E_iV_i\rangle\big) \\ &= \operatorname*{argmin}_{V_i \in \mathbb{C}^{r_{i-1} \times q_i \times r_i}} \big(\frac{1}{2}\langle E_i^\dagger \mathbf{X}E_iV_i, V_i\rangle - \langle E_i^\dagger B, V_i\rangle\big) \end{aligned} \tag{136}$$

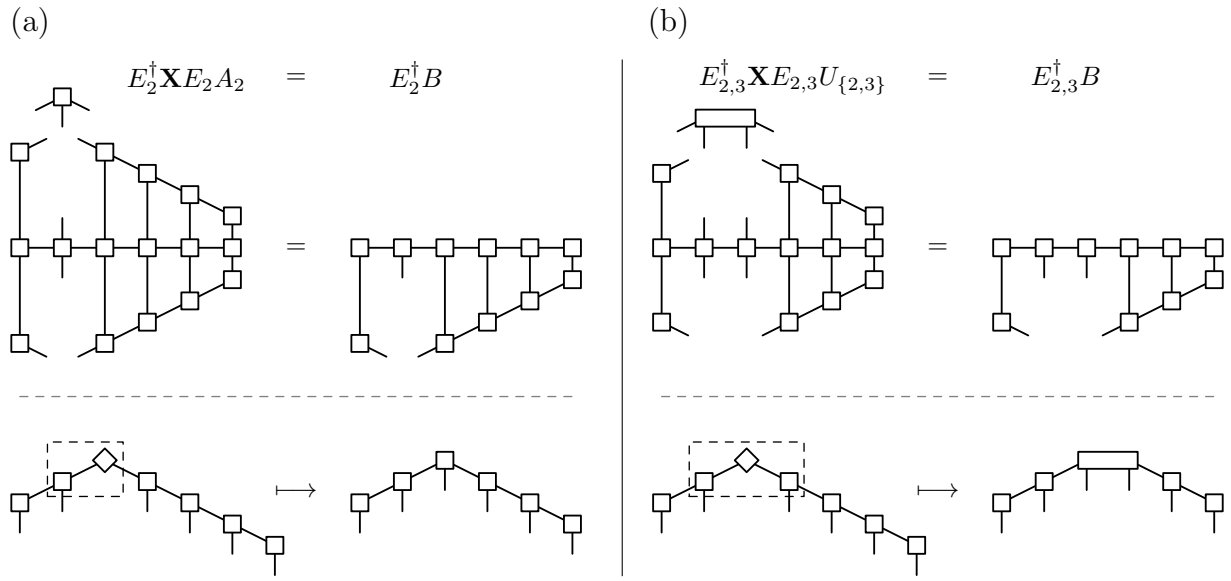


Figure 11: A micro-iteration of (a) the ALS algorithm, and (b) the two-site DMRG.

At stationary points V_i of the functional $j \circ E_i$, there holds the first order condition

$$\nabla(j \circ E_i)(V_i) = E_i^\dagger \mathbf{X} E_i \mathcal{V}_i - E_i^\dagger B = 0. \tag{137}$$

As such, one micro-iteration of the ALS algorithm can be defined as

$$U^+ := A_1 \circ \cdots \circ A_i^+ \circ \cdots \circ A_d \tag{138}$$

$$A_i^+ = \left(E_i^\dagger \mathbf{X} E_i\right)^{-1} E_i^\dagger B. \tag{139}$$

See Fig. 11(a) for an illustration.

In the subspace notation, we get

$$|\Psi\rangle^+ = \operatorname*{argmin}_{|\Phi\rangle \in \mathcal{L}_i \otimes \Lambda_i \otimes \mathcal{R}_i} \left(\frac{1}{2}\langle \Phi|\mathbf{X}|\Phi\rangle - \langle \Upsilon, \Phi\rangle\right) = P_i \mathbf{X}^{-1} P_i |\Upsilon\rangle. \tag{140}$$

For this to work, $\mathbf{X}$ does not necessarily have to be invertible on the whole tensor space but only on the small subspaces $\mathcal{L}_i \otimes \Lambda_i \otimes \mathcal{R}_i$. This is guaranteed if $\mathbf{X}$ is invertible as a whole. Additionally, one can see that the spectrum of $\mathbf{X}$ on $\mathcal{L}_i \otimes \Lambda_i \otimes \mathcal{R}_i$ is a subset of the whole spectrum of $\mathbf{X}$ and in particular it holds $\mathrm{cond}_i(\mathbf{X}) \leq \mathrm{cond}(\mathbf{X})$[99].

This notation suggests that the ALS is closely related to the DMRG algorithm. In fact, it is often called the *one-site* DMRG as it can be seen as a simple modification of the that algorithm. In comparison, the ALS has the advantage that it optimizes the tensor on very small subspaces. On the other hand, the ranks $\mathbf{r} = (r_1, \ldots, r_{d-1})$ remain fixed and have to be guessed at the beginning. In order to introduce higher ranks, one has to do this in a greedy fashion, e.g. by adding a rank-one approximation of the residual, see[233].

The classical two-site DMRG is a clever modification. Here, we minimize over the bigger subspace $\mathcal{L}_i \otimes \Lambda_i \otimes \Lambda_{i+1} \otimes R_{i+1}$, with the basis representation as in Fig. 10(b),

$$|\Psi\rangle = \sum_{m_{i-1}, m_{i+1}, \alpha_i, \alpha_{i+1}} U_{\{i,i+1\}}(m_{i-1}, \alpha_i, \alpha_{i+1}, m_{i+1}) \; |\xi_{m_{i-1}}^{\{i-1\}}\rangle \otimes |\phi_{\alpha_i}^{\{i\}}\rangle \otimes |\phi_{\alpha_{i+1}}^{\{i+1\}}\rangle \otimes \zeta_{m_{i+1}}^{\{i+2\}}\rangle, \tag{141}$$

	canonical	Tucker	TT	HT
complexity	$\mathcal{O}(qdr)$	$\mathcal{O}(r^d + qdr)$	$\mathcal{O}(qdr^2)$	$\mathcal{O}(qdr + dr^3)$
rank	no	defined	defined	defined
	$\max\{\mathbf{r}_{Tucker}\} \leq \max\{\mathbf{r}_{HT}\} \leq r_{canonical}$			
(weak) closedness	no	yes	yes	yes
ALS (1site DMRG)	yes - but slow	yes	yes	yes
DMRG	no	no	yes	no
H(O)SVD	no	yes	yes	yes
embedded manifold	no	yes	yes	yes
dyn. low rank approx.	no	yes	yes	yes
algebraic variety	no	yes	yes	yes
recovery	??	yes	yes	yes
quasi best approx.	no	yes	yes	yes
best approx.	no	exist but NP-hard	exist but NP-hard	exist but NP-hard

Table 1: Comparison between the different tensor formats introduced in Sec. 3.

i.e. we optimize two components at the same time, see Fig. 11(b). The advantage is that a subsequent SVD after the optimization step in order to separate the two components yields a new - and possibly higher - rank. To control the size of these new ranks, a further truncation is often required. Several strategies for *dynamical rank selection* can be implemented by considering the error in different norms[99,125].

General convergence theory of both the ALS and the DMRG is subject to research[203]. They converge only to stationary points or at most local minima, as global convergence can not be guaranteed[99]. Some convergence results have been published for a modified scheme, that proceeds in a Gauß-Southwell-like fashion and optimizes only the component with the largest residual[138,210]. There are also many open questions in dealing with physical applications like the SCF iteration for Hartree Fock models. The prescribed approach is completely variational, which has important consequences for computing gradients, e.g. forces. An efficient implementation plays a crucial role. The interested reader should consult fundamental contributions in the DMRG literature[39,127,249,250].

As a summary of the section about tensor formats, we present Tab. 1 that compares the different decompositions, their complexity and their advantages and disadvantages in numerical computations.

4 Numerical techniques

In order to utilize efficiently the theoretical framework discussed in the previous sections one has to carry out various optimization tasks. Therefore, we start this section with a brief overview and highlight important concepts using simple examples. Then various iterative methods based on blocking procedure will be reviewed briefly and the concept of entanglement will be studied with respect to entanglement localization, geometrical network optimization, choosing optimal bases and network initialization. In this section, our focus is

on the numerical method, thus entropic measures of electronic properties of molecules will be discussed only very briefly.

For pedagogical reasons, tutorial examples will be presented for a text book example, the LiF molecule. Due to the ionic-neutral curve crossing between the two lowest $^1\Sigma^+$ states of LiF, this system provides a good testing ground to demonstrate the efficiency of the *quantum chemistry version of the density matrix renormalization group* method (QC-DMRG) and *tree tensor network state* (QC-TTNS) algorithm. Our analysis is especially useful for systems in which the wave function character of molecules changes as a function of geometry. In the LiF example, it differs greatly on two sides of an avoided crossing in a diatomic molecule. Atomic orbital (AO) basis was adapted from the literature[19] in order to match with previous DMRG computations[126]. The AO basis set[19] is suitable to describe the ionic and covalent LiF states as well. It consists of 9s and 4p functions contracted to 4s and 2p functions on the Li atom and 9s, 6p and 1d functions contracted to 4s, 3p and 1d on the F atom. For more details of the AO basis set, we refer to the original publication[19]. The two lowest $^1\Sigma^+$ states of LiF around the equilibrium bond length can be qualitatively described by the $1\sigma^2 2\sigma^2 3\sigma^2 4\sigma^2 1\pi^4$ and $1\sigma^2 2\sigma^2 3\sigma^2 4\sigma^1 5\sigma^1 1\pi^4$ configurations[75]. For this reason, the MO basis was obtained by CASSCF optimizations, with two active electrons on two active orbitals (4σ and 5σ) (CAS(2,2)). MO's were optimized simultaneously for both $^1\Sigma^+$ states. T_{ij} and V_{ijkl} matrix elements of Eq. (44a) and (44b) are expressed in this MO basis. CASSCF optimizations were carried out with the `GAMESS-US` quantum chemistry package[75]. Orbitals 1σ, 2σ and 3σ were kept frozen in all presented configurational interaction (CI), MPS(DMRG) and TTNS computations. Six of the valence electrons were excited to all orbitals in the CI calculation, which we use as reference to benchmark the QC-DMRG and QC-TTNS results. Therefore the active space in most of our CI, MPS(DMRG) and TTNS computations consists of 6 electrons and 25 orbitals: CAS(6,25). In certain cases, a smaller active space, CAS(6,12), will also be used. Using the same MO basis obtained as a result of CASSCF optimizations in the previous CAS(2,2) active space, the CAS(6,12) active space is constructed by excluding the three lowest lying occupied and 13 highest virtual orbitals from the total 28 orbitals. CI results were obtained by utilizing standard full-CI programs. C_{2v} point group symmetry constraints were assigned during this study.

4.1 Basic terms and brief overview

4.1.1 The problem in the language of tensor factorization

Let us start this section with a very brief summary in order to highlight the most important concepts. In the rest of the paper, a spin-orbital will be called a *local tensor space* $\Lambda \cong \mathbb{C}^q$, with $\dim \Lambda = q$, and will be denoted by $\bullet$. Using the fermionic occupation number basis ($q = 2$), $|\mu, s\rangle$ for all spins $s \in \{\downarrow, \uparrow\}$, with $\mu \in \{0, 1\}$ occupation numbers, the operators (see (35)) are defined as

$$\mathbf{a}^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{s} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{142}$$

where $\mathbf{a}^\dagger$ creates an electron, $\mathbf{I}$ is the identity matrix and $\mathbf{s}$ stands for the phase factor due to the antisymmetric fermionic wavefunction. As was constructed in section 2.7, it is

$\alpha_{\{1,2\}}$	α_1	α_2	$N_{\alpha_{\{1,2\}}\uparrow}$	$N_{\alpha_{\{1,2\}}\downarrow}$	$N_{\alpha_1\uparrow}$	$N_{\alpha_1\downarrow}$	$N_{\alpha_2\uparrow}$	$N_{\alpha_2\downarrow}$
1	$-$	$-$	0	0	0	0	0	0
2	$-$	$\downarrow$	0	1	0	0	0	1
3	$-$	$\uparrow$	1	0	0	0	1	0
4	$-$	$\uparrow\downarrow$	1	1	0	0	1	1
5	$\downarrow$	$-$	0	1	0	1	0	0
6	$\downarrow$	$\downarrow$	0	2	0	1	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
16	$\uparrow\downarrow$	$\uparrow\downarrow$	2	2	1	1	1	1

Table 2: Basis states for a two-orbital system. Index values of basis states are $\alpha_1, \alpha_2 \in \{1,2,3,4\}$, and we use the shorthand notation $|\phi^{\{1\}}_{\alpha_1}\rangle, |\phi^{\{2\}}_{\alpha_2}\rangle \in \{|\phi_1\rangle \equiv |-\rangle, |\phi_2\rangle \equiv |\downarrow\rangle, |\phi_3\rangle \equiv |\uparrow\rangle, |\phi_4\rangle \equiv |\uparrow\downarrow\rangle\}$, as usual. For the two-site basis $\alpha_{\{1,2\}} = (\alpha_1 - 1)q + \alpha_2 \in \{1,2,3,4\ldots 16\}$, and $|\phi^{\{1,2\}}_{\alpha_{\{1,2\}}}\rangle \in \{|\phi^{\{1,2\}}_1\rangle \equiv |--\rangle, |\phi^{\{1,2\}}_2\rangle \equiv |-\downarrow\rangle, |\phi^{\{1,2\}}_3\rangle \equiv |-\uparrow\rangle, |\phi^{\{1,2\}}_4\rangle \equiv |-\uparrow\downarrow\rangle, \ldots, |\phi^{\{1,2\}}_{16}\rangle \equiv |\uparrow\downarrow\uparrow\downarrow\rangle\}$. Particle numbers for different spins are also shown. These are proper quantum numbers if the corresponding operators commute with the Hamiltonian.

also possible to use a $\mathbb{C}^4$ representation in which case $\bullet$ will represent a molecular orbital ($q = 4$). In this representation a state can be empty, singly occupied with spin-up or down particle, or doubly occupied, represented by the basis states $\{|\phi_\alpha\rangle\}$ for $\alpha \in \{1,2,3,4\}$ as $\{|\phi_1\rangle \equiv |-\rangle, |\phi_2\rangle \equiv |\downarrow\rangle, |\phi_3\rangle \equiv |\uparrow\rangle, |\phi_4\rangle \equiv |\uparrow\downarrow\rangle\}$. (In this sloppy but extremely convenient notation, on the one hand, $|\phi_\alpha\rangle \equiv |\alpha\rangle$ is written for simplicity, usual in quantum information theory, on the other hand, the $1,2,3,4$ index-values (useful for computers) are identified with the $-,\downarrow,\uparrow,\uparrow\downarrow$ labels of the states (carrying physical meaning). Therefore we can write the same basis state in four different ways, e.g., $|\phi_2\rangle \equiv |\phi_\downarrow\rangle \equiv |2\rangle \equiv |\downarrow\rangle$.) The relevant orbital operators (51a)-(51b) in this basis are

$$\mathbf{c}^\dagger_\uparrow = \mathbf{a}^\dagger \otimes \mathbf{I} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad \mathbf{c}^\dagger_\downarrow = \mathbf{s} \otimes \mathbf{a}^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \tag{143a}$$

$$\mathbb{I} = \mathbf{I} \otimes \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{z} = \mathbf{s} \otimes \mathbf{s} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{143b}$$

We can put together two $\mathbb{C}^4$ tensor spaces, i.e., forming a two-orbital system ($\bullet\bullet$), where $\Lambda^{\{1,2\}} = \Lambda_1 \otimes \Lambda_2$ with $\dim \Lambda^{\{1,2\}} = \dim \Lambda_1 \dim \Lambda_2 = q^2 = 16$. The basis of the $\bullet\bullet$ system is given as $|\phi^{\{1,2\}}_{\alpha_{\{1,2\}}}\rangle = |\phi^{\{1\}}_{\alpha_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle$ where $\alpha_{\{1,2\}} = (\alpha_1 - 1)q + \alpha_2$. The relevant operators for the $\bullet\bullet$ system are formed as

$$\mathbf{c}^\dagger_{1,\uparrow} = \mathbf{c}^\dagger_\uparrow \otimes \mathbb{I}, \qquad \mathbf{c}^\dagger_{2,\uparrow} = \mathbf{z} \otimes \mathbf{c}^\dagger_\uparrow, \qquad \mathbf{c}^\dagger_{1,\downarrow} = \mathbf{c}^\dagger_\downarrow \otimes \mathbb{I}, \qquad \mathbf{c}^\dagger_{2,\downarrow} = \mathbf{z} \otimes \mathbf{c}^\dagger_\downarrow. \tag{144}$$

A wavefunction (33a), (53) can be expressed in a general form as

$$|\Psi^{\{1,2\}}\rangle = \sum_{\alpha_1,\alpha_2} U^{\{1,2\}}(\alpha_1,\alpha_2) |\phi^{\{1\}}_{\alpha_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle, \tag{145}$$

where the matrix $U^{\{1,2\}}(\alpha_1,\alpha_2)$ describes the quantum mechanical probability distrubution of the basis of the combined system. Such wavefunctions can arise from the diagonalization of the Hamiltonian $\mathbf{H}$, which is a q^2 by q^2 matrix (43), using the above representation of the creation and annihilation operators. The full diagonalization of $\mathbf{H}$ gives the exact solution (full-CI), and the m^{th} eigenstate of a two-orbital Hamiltonian is

$$|\Psi_m^{\{1,2\}}\rangle = \sum_{\alpha_1,\alpha_2} U^{\{1,2\}}(\alpha_1,\alpha_2,m)|\phi_{\alpha_1}^{\{1\}}\rangle \otimes |\phi_{\alpha_2}^{\{2\}}\rangle, \tag{146}$$

where $\alpha_1,\alpha_2 = 1,\ldots,q$ and $m = 1,\ldots,q^2$.

4.1.2 Change of basis, truncation and iterative diagonalization

The representation of the problem is, however, not unique. Using a unitary operator acting on $\Lambda^{\{1,2\}}$, $\mathbf{O}$, which leaves the eigenvalue spectrum of the Hamiltonian unchanged, we can carry out a *change of basis*

$$\mathbf{H} \qquad \longmapsto \qquad \mathbf{OHO}^\dagger. \tag{147}$$

One possibility to achive this is to apply the unitary operator to all operators used to construct the Hamiltonian 43, i.e., $\mathbf{c}_{1,\downarrow}^\dagger \mapsto \mathbf{O}\mathbf{c}_{1,\downarrow}^\dagger\mathbf{O}^\dagger$, $\mathbf{c}_{2,\downarrow}^\dagger \mapsto \mathbf{O}\mathbf{c}_{2,\downarrow}^\dagger\mathbf{O}^\dagger$, $\mathbb{I} \mapsto \mathbf{O}\mathbb{I}\mathbf{O}^\dagger = \mathbb{I}$, $\mathbf{c}_{1,\uparrow}^\dagger\mathbf{c}_{2,\uparrow} \mapsto \mathbf{O}\mathbf{c}_{1,\uparrow}^\dagger\mathbf{c}_{2,\uparrow}\mathbf{O}^\dagger$, etc. If the rows of the matrix $\mathbf{O}$ is constructed from the $|\Psi_m^{\{1,2\}}\rangle$ eigenstates (146) then we arrive at the eigenbasis representation of $\mathbf{H}$, i.e., $\mathbf{H}$ becomes diagonal and its elements are equivalent to the eigenvalues of the original problem.

The eigenvalue spectrum of $\mathbf{H}$ determines the physical properties of the system exactly. It is, however, possible to use an approximate representation of $\mathbf{H}$, i.e., using a smaller basis as we select only $M < q^2$ eigenstates to form the $\mathbf{O}$ matrix, which becomes then rectangular. That is, we change over to a subspace $\Xi^{\{1,2\}}$ of the original tensor space $\Lambda^{\{1,2\}}$, (see section 3). This *truncation* leads to loss of information as $\mathbf{OO}^\dagger \neq \mathbb{I}$, but the kept eigenstates can still provide a good description of the low-energy physics of the problem.

If we are interested in the low-lying eigenstates of $\mathbf{H}$ it is not necessary to carry out a full diagonalization, but systematic application of the Hamiltonian to a randomly chosen state provides the lowest lying eigenstate. An extension of such power methods, like the *Lánczos*[118] or *Davidson*[55] methods, provides faster convergence rates, and excited states can also be calculated[93,178,211].

4.1.3 Unitary transformation for two molecular orbitals

For the sake of simplicity let us consider an example of two $S = 1/2$-spins. That is, the basis of the $\bullet\bullet$ system is formed from the $|\phi_1\rangle \equiv |\downarrow\rangle$, $|\phi_2\rangle \equiv |\uparrow\rangle$ vectors with $q = 2$, and the eigenvectors (146) can be formed as

$$|\Psi_m^{\{1,2\}}\rangle = \sum_{\alpha_1,\alpha_2} O(m, 2(\alpha_1-1)+\alpha_2)|\phi_{\alpha_1}^{\{1\}}\rangle \otimes |\phi_{\alpha_2}^{\{2\}}\rangle, \tag{148}$$

where $\alpha_1,\alpha_2 \in \{1,2\} \equiv \{\downarrow,\uparrow\}$. An example for the $\mathbf{O}$ matrix is shown in Table 3. The dimension of the O matrix is $M \times q^2$ where M can take values between 1 and q^2 (truncation). The Mq^2 elements of the matrix can also be represented by q (two) $M \times q$ matrices, denoted

O	$\downarrow\downarrow$	$\downarrow\uparrow$	$\uparrow\downarrow$	$\uparrow\uparrow$	S^z	S
$\Psi_1^{\{1,2\}}$	1	0	0	0	-1	1
$\Psi_2^{\{1,2\}}$	0	$1/\sqrt{2}$	$1/\sqrt{2}$	0	0	1
$\Psi_3^{\{1,2\}}$	0	$1/\sqrt{2}$	$-1/\sqrt{2}$	0	0	0
$\Psi_4^{\{1,2\}}$	0	0	0	1	1	1

Table 3: An example for the unitary matrix $\mathbf{O}$, used to transfrom the Hamilton to an S^z eigenbasis. This transformation arises when the Hamiltonian commutes with the operators of the z-component and the magnitude of the total spin. Then the eigenvalues of these operators, S^z and S respectively, are proper quantum numbers, and are listed in the last two columns of the table.

with $\mathbf{B}_2(\alpha_2)$, i.e., for each basis of the second spin we assign a matrix. This means that we take columns 1 and 3 to form $\mathbf{B}_2(\downarrow)$ and columns 2 and 4 for $\mathbf{B}_2(\uparrow)$, so, for the example given in Table 3 we have (without truncation)

$$(\mathbf{B}_2(\downarrow))_{m,\alpha_1} = \begin{pmatrix} 1 & 0 \\ 0 & 1/\sqrt{2} \\ 0 & -1/\sqrt{2} \\ 0 & 0 \end{pmatrix}, \qquad (\mathbf{B}_2(\uparrow))_{m,\alpha_1} = \begin{pmatrix} 0 & 0 \\ 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}. \tag{149}$$

We also denote this by $(\mathbf{B}_2(\alpha_2))_{m,\alpha_1} = B_2(m, \alpha_2, \alpha_1)$. It is easy to recognize that such $\mathbf{B}$ matrices form the basis of the matrix product state representation discussed in Sec.3. In the literature, usually $\mathbf{A}_2 \equiv \mathbf{B}_2^T$ is used, that is, $A_2(\alpha_1, \alpha_2, m) = B_2(m, \alpha_2, \alpha_1)$, and the wavefunction is written as

$$|\Psi_m^{\{1,2\}}\rangle = \sum_{\alpha_1,\alpha_2} A_2(\alpha_1, \alpha_2, m) |\phi_{\alpha_1}^{\{1\}}\rangle \otimes |\phi_{\alpha_2}^{\{2\}}\rangle. \tag{150}$$

4.1.4 Symmetries

In many systems the time evolution governed by the Hamilton operator does not change the value of a measurable quantity, i.e., the Hamilton operator commutes with the operator associated to that measurable quantity. These operators are called symmetry operators and can be used to cast the Hilbert space to smaller independent subspaces. Consequently, instead of solving a large matrix eigenvalue problem, the eigenvalue spectrum can be determined by solving several smaller problems. Thus, the distinct quantum numbers helps to partition the Hilbert space into multiple independent subspaces corresponding to a given combination of quantum number values.

A given symmetry operator has the same eigenvectors as the Hamiltonian, thus the eigenstates of the Hamiltonian can be labelled by the eigenvalues of the symmetry operator (*quantum number* Q), and the Hilbert space can be decomposed into subspaces (*sectors*) spanned by the eigenvectors of each quantum number value[52]. Introducing a quantum number based representation, the sparse operators can be decomposed to a set of smaller but dense matrices, furthermore the Hamiltonian operator becomes blockdiagonal.

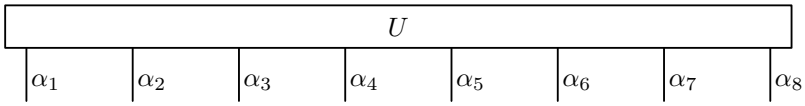


Figure 12: Example $d = 8$.

For two orbitals, quantum numbers are formed from orbital quantum numbers as $Q_{\{\alpha_i,\alpha_j\}} = f(Q_{\alpha_i}, Q_{\alpha_j})$, where function f depends on the given symmetry. For $U(1)$ symmetries the $f(Q_{\alpha_i}, Q_{\alpha_j}) = Q_{\alpha_i} + Q_{\alpha_j}$ while for non-Abelian symmeties, such as for the conservation of total spin, more complex algebra is involved, based on the Wigner-Eckart theorem [231,247,254–256]. For more details, see section 4.4.8.

4.1.5 Unitary transformation for d number of molecular orbitals and tensor product approximation

The formalism discussed above can be extended to describe a system with d molecular orbitals denoted as $\bullet\bullet\bullet\ldots\bullet$. The Hilbert space is formed as $\Lambda^{\{1,2,\ldots,d\}} = \otimes_{i=1}^{d}\Lambda_i$ with $\dim\Lambda^{\{1,2,\ldots,d\}} = \prod_{i=1}^{d}\dim\Lambda_i = q^d$. A wavefunction is written as

$$|\Psi^{\{1,2,\ldots,d\}}\rangle = \sum_{\alpha_1\ldots\alpha_d} U^{\{1,2,\ldots,d\}}(\alpha_1,\alpha_2,\ldots\alpha_d)|\phi^{\{1\}}_{\alpha_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle \otimes \ldots \otimes |\phi^{\{d\}}_{\alpha_d}\rangle, \tag{151}$$

where $U^{\{1,2,\ldots,d\}}$ is a tensor of order d. Since the dimension of U scales exponentially with d we need approximative methods. The major aim is to find a good approximation of U in terms of products of lower order tensors with smaller rank than the original problem.

One possibility is to systematically apply the procedure outlined in Sec. 4.1.2 and in Sec. 3 to describe one, two, three, ... d-orbital wavefunction. Starting with two orbitals, $\bullet\bullet$, the new (truncated) basis of the composed system is written as

$$|\xi^{\{1,2\}}_{m_2}\rangle = \sum_{\alpha_1,\alpha_2} A_2(\alpha_1,\alpha_2,m_2)|\phi^{\{1\}}_{\alpha_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle, \tag{152a}$$

which is shown schematically in Fig. 13(a). This can be rewritten as

$$|\xi^{\{1,2\}}_{m_2}\rangle = \sum_{m_1,\alpha_2} A_2(m_1,\alpha_1,m_2)|\xi^{\{1\}}_{m_1}\rangle \otimes |\phi^{\{2\}}_{\alpha_2}\rangle, \tag{152b}$$

using the identity

$$|\xi^{\{1\}}_{m_1}\rangle = \sum_{\alpha_1} A_1(1,\alpha_1,m_1)|\phi^{\{1\}}_{\alpha_1}\rangle \quad \text{with} \quad A_1(1,\alpha_1,m_1) = \delta_{\alpha_1,m_1}, \tag{152c}$$

as is depicted in Fig. 13(b). We have the above form for $A_1(1,\alpha_1,m_1)$ since here the transformation and truncation comes from the subspace approximation in $\Lambda^{\{1,2\}}$ common in a wide part of renormalization group methods in physics, shown in sections 4.1.2 and 4.1.3. On the other hand, when the transformations and truncation come from successive subspace

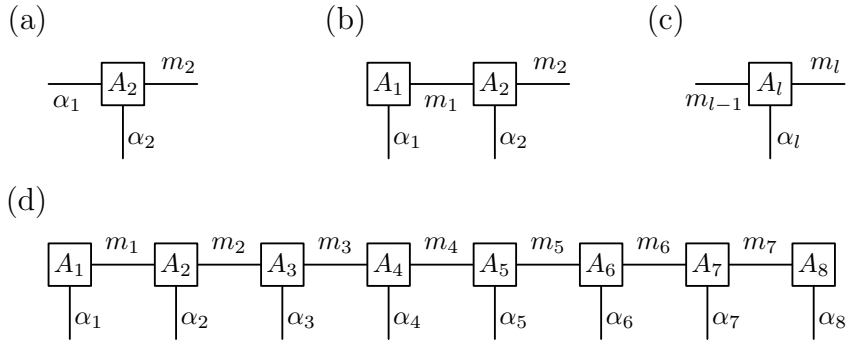


Figure 13: (a) Graphical representation of the two-orbital composed system using the procedure outlined in Sec. 4.1.2, (b) and after using idenity (152c). (c) Graphical representation of the component tensor and (d) the $d = 8$-orbital wavefuntion as a network built from matrices.

optimization starting with the space Λ_1, e.g., based on SVD, then we have nontrivial basis change even inside Λ_1, see section 3.5. For three orbitals $\bullet\bullet\bullet$,

$$|\xi_{m_3}^{\{1,2,3\}}\rangle = \sum_{m_2,\alpha_3} A_3(m_2,\alpha_3,m_3)|\xi_{m_2}^{\{1,2\}}\rangle \otimes |\phi_{\alpha_3}^{\{3\}}\rangle. \tag{152d}$$

This procedure can be extended iteratively using series of component tensors

$$|\xi_{m_l}^{\{1,2,\ldots,l\}}\rangle = \sum_{\alpha_1,\ldots\alpha_l} (\mathbf{A}_1(\alpha_1)\mathbf{A}_2(\alpha_2)\ldots\mathbf{A}_l(\alpha_l))_{1,m_l}|\phi_{\alpha_1}^{\{1\}}\rangle \otimes |\phi_{\alpha_2}^{\{2\}}\rangle \otimes \ldots \otimes |\phi_{\alpha_l}^{\{l\}}\rangle, \tag{152e}$$

where the component tensor $A_l(m_{l-1},\alpha_l,m_l) = (\mathbf{A}_l(\alpha_l))_{m_{l-1},m_l}$ is defined as

$$|\xi_{m_l}^{\{1,2,\ldots,l\}}\rangle = \sum_{m_{l-1},\alpha_l} (\mathbf{A}_l(\alpha_l))_{m_{l-1},m_l}|\xi_{m_{l-1}}^{\{1,2,\ldots,l-1\}}\rangle \otimes |\phi_{\alpha_l}^{\{l\}}\rangle, \tag{152f}$$

see in Fig. 13(c). As a result of this procedure, the d-orbital wavefunction is expressed as

$$|\Psi\rangle = \sum_{\alpha_1,\alpha_2,\ldots,\alpha_d} \mathbf{A}_1(\alpha_1)\mathbf{A}_2(\alpha_2)\cdots\mathbf{A}_d(\alpha_d)\ |\phi_{\alpha_1}^{\{1\}}\rangle \otimes |\phi_{\alpha_2}^{\{2\}}\rangle \otimes \ldots \otimes |\phi_{\alpha_d}^{\{d\}}\rangle, \tag{152g}$$

i.e., for each molecular orbital we can assign a matrix $\mathbf{A}_l(\alpha_l)$, coming from the basis change in $\Lambda^{(l-1)} \otimes \Lambda_l$, and we form a network built from matrices as shown in Fig. 13(d). For more detailed derivations we refer to the original papers and review articles[67,81,184,212,236,243] and Sec. 3.

Succesively repeating the construction of section 4.1.4, the quantum numbers for the q^d states of the d-orbital systems can be determined. As before, the full Hilbert space is decomposed into sectors based on these quantum numbers. If we consider only the case where the number of electrons with down and up spins is conserved, the quantum number is the vector $Q = (N_\downarrow, N_\uparrow)$ with $N = N_\downarrow + N_\uparrow$, then the dimension of the related sector $\Lambda_{\mathrm{FCI}} \subset \Lambda^{(d)}$ in the Hilbert space is $\dim \Lambda_{\mathrm{FCI}} = \binom{d}{N_\downarrow}\binom{d}{N_\uparrow}$.

4.1.6 Tensor topology

If we render the tensor spaces corresponding to the orbitals in a "one- or two-dimensional space" (higher dimensional extension is also possible) we form a *chain-* or *lattice-topology* of the tensor product representation. In some cases this topology is also reflected by the physical lattice topology of the problem, i.e., one-dimensional-like polimers can be studied very well using the one-dimensional tensor-topology. As will be discussed below, one of the major aim is to find the best tensor topology for a given molecule.

4.2 Entanglement and correlations

In the previous subsection, we have considered basis change based on the Hamiltonian of the system. Another approach of basis change is based on an actual pure state of the system, and connected to the *entanglement* of that state[103,227].

In quantum systems, correlations having no counterpart in classical physics arise. Pure states showing these strange kinds of correlations are called entangled[103,227], and the existence of these states has so deep and important consequences[21,48,63] that Schrödinger has identified entanglement to be the characteristic trait of quantum mechanics[214,215]. The QC-DMRG and QC-TTNS algorithms approximate a composite system with long-range interactions, and it turned out that the results of quantum information theory[177,257] can be used to understand the criteria of its convergence.

4.2.1 Singular value decomposition and entanglement

The basic concept on which entanglement theory is built up is the entanglement with respect to a *bipartition* of the system. In this manybody situation, the system composed of d orbitals can be treated as the sum of two subsystems (also called blocks), $(\mathrm{A}), (\mathrm{B}) \subset \{1, 2, \ldots, d\}$. (They are disjoint, and their union gives the whole system.) The Hilbert spaces associated to them are $\Lambda^{(\mathrm{A})}$ and $\Lambda^{(\mathrm{B})}$, so $\Lambda^{\{1,2,\ldots,d\}} \cong \Lambda^{(\mathrm{A})} \otimes \Lambda^{(\mathrm{B})}$. After choosing bases in the subsystems, $\{|\phi^{(\mathrm{A})}_{\alpha_{(\mathrm{A})}}\rangle \in \Lambda^{(\mathrm{A})}\}$ and $\{|\phi^{(\mathrm{B})}_{\alpha_{(\mathrm{B})}}\rangle \in \Lambda^{(\mathrm{B})}\}$, the wavefunction (151) characterizing the *pure state* of the system can be written as

$$|\Psi\rangle = \sum_{\alpha_{(\mathrm{A})}=1}^{\dim \Lambda^{(\mathrm{A})}} \sum_{\alpha_{(\mathrm{B})}=1}^{\dim \Lambda^{(\mathrm{B})}} U(\alpha_{(\mathrm{A})}, \alpha_{(\mathrm{B})}) |\phi^{(\mathrm{A})}_{\alpha_{(\mathrm{A})}}\rangle \otimes |\phi^{(\mathrm{B})}_{\alpha_{(\mathrm{B})}}\rangle. \tag{153}$$

Based on the UDV-decomposition of the matrix $U(\alpha_{(\mathrm{A})}, \alpha_{(\mathrm{B})})$, one can find a product unitary transformation $\mathbf{O}_A \otimes \mathbf{O}_B$, which brings it to the *Schmidt form*[209]

$$|\Psi\rangle = \sum_{m=1}^{r_{\mathrm{Sch}}} \sqrt{\omega_m} |\xi^{(\mathrm{A})}_m\rangle \otimes |\xi^{(\mathrm{B})}_m\rangle. \tag{154}$$

Here, the vectors $|\xi^{(\mathrm{A})}_{m_{(\mathrm{A})}}\rangle$ and $|\xi^{(\mathrm{B})}_{m_{(\mathrm{B})}}\rangle$ form orthonormal bases, also called *Schmidt bases*, in the Hilbert spaces of the two blocks, $\langle \xi^{(\mathrm{A})}_m | \xi^{(\mathrm{A})}_{m'} \rangle = \langle \xi^{(\mathrm{B})}_m | \xi^{(\mathrm{B})}_{m'} \rangle = \delta_{m,m'}$, moreover, the squares of the *Schmidt coefficients* $\sqrt{\omega_m}$ satisfy $0 \leq \omega_m \leq 1$ with the constraint $\sum_m \omega_m = 1$. The summation goes until the *Schmidt rank*, $r_{\mathrm{Sch}} \leq \min(\dim \Lambda^{(\mathrm{A})}, \Lambda^{(\mathrm{B})})$. The $\sqrt{\omega_m}$

numbers are also called the singular values of $U(\alpha_{(\mathrm{A})}, \alpha_{(\mathrm{B})})$, and the above form *singlar value decomposition* (SVD). If the *Schmidt rank* $r_{\mathrm{Sch}} > 1$, then $|\Psi\rangle$ is entangled (inseparable) and we say that the two blocks are entangled[103].

If we consider the two-electron subspace of a two-orbital system, then the state

$$|\Psi_{\mathrm{ent}}\rangle = \frac{1}{\sqrt{2}}\big(|\phi_1^{\{1\}}\rangle \otimes |\phi_2^{\{2\}}\rangle - |\phi_2^{\{1\}}\rangle \otimes |\phi_1^{\{2\}}\rangle\big) \equiv \frac{1}{\sqrt{2}}\big(|\downarrow\rangle \otimes |\uparrow\rangle - |\uparrow\rangle \otimes |\downarrow\rangle\big) \tag{155a}$$

is an entangled state, while

$$|\Psi_{\mathrm{sep}}\rangle = |\phi_1^{\{1\}}\rangle \otimes |\phi_2^{\{2\}}\rangle \equiv |\downarrow\rangle \otimes |\uparrow\rangle \tag{155b}$$

is separable. Both vectors are almost in Schmidt form, (unitary transformation $\mathbf{O}_2$ acting as $|\phi_2^{\{2\}}\rangle \mapsto |\phi_1^{\{2\}}\rangle$ and $|\phi_1^{\{2\}}\rangle \mapsto -|\phi_2^{\{2\}}\rangle$ brings the first one to a Schmidt form) and the squared Schmidt coefficients can immediately be read: $\omega_1 = \omega_2 = 1/2$ in the first case and $\omega_1 = 1$, $\omega_2 = 0$ in the second.

For a system characterized by a pure state $|\Psi\rangle$, the *state of the subsystem* (A) is encoded in the *reduced density matrix* of the subsystem,

$$\rho^{(\mathrm{A})} = \mathrm{Tr}_{\mathrm{B}}\, |\Psi\rangle\langle\Psi|. \tag{156}$$

The subsystem of interest is usually labelled by (A) and the other subsystem (B), which can also be considered as the "environment" of (A). The operation Tr_{B} means carrying out the trace over subsystem (B), that is, $\mathrm{Tr}_{\mathrm{B}}(\mathbf{X} \otimes \mathbf{Y}) = \mathbf{X}\mathrm{Tr}(\mathbf{Y})$, leading to the form

$$\rho^{(\mathrm{A})} = \sum_{\alpha_{(\mathrm{A})}, \alpha'_{(\mathrm{A})}} \left[\sum_{\alpha_{(\mathrm{B})}} U(\alpha_{(\mathrm{A})}, \alpha_{(\mathrm{B})})\overline{U(\alpha'_{(\mathrm{A})}, \alpha_{(\mathrm{B})})}\right] |\phi^{(\mathrm{A})}_{\alpha_{(\mathrm{A})}}\rangle\langle\phi^{(\mathrm{A})}_{\alpha'_{(\mathrm{A})}}|, \tag{157}$$

having the matrix elements in the square bracket

$$\rho^{(\mathrm{A})}(\alpha_{(\mathrm{A})}, \alpha'_{(\mathrm{A})}) \equiv \langle\phi^{(\mathrm{A})}_{\alpha_{(\mathrm{A})}}|\rho^{(\mathrm{A})}|\phi^{(\mathrm{A})}_{\alpha'_{(\mathrm{A})}}\rangle = \sum_{\alpha_{(\mathrm{B})}} U(\alpha_{(\mathrm{A})}, \alpha_{(\mathrm{B})})\overline{U(\alpha'_{(\mathrm{A})}, \alpha_{(\mathrm{B})})} \tag{158}$$

(Similar expressions can be written for subsystem (B).) If we write (156) using the Schmidt form (154), we get immediately a diagonal form

$$\rho^{(\mathrm{A})} = \sum_m \omega_m |\xi^{(\mathrm{A})}_m\rangle\langle\xi^{(\mathrm{A})}_m| \tag{159}$$

in the Schmidt basis.

On the other hand, in this pure case, the information on the *entanglement* between the (A) and (B) blocks of the system is encoded in the density matrices of the blocks. It turns out that the eigenvalue spectrum of $\rho^{(\mathrm{A})}$ is enough for the complete characterization of the entanglement between blocks (A) and (B), and, as we have seen in (159), it follows from the Schmidt decomposition that the eigenvalues of $\rho^{(\mathrm{A})}$ are exactly the squared Schmidt coefficients ω_m in (154). (The same holds for $\rho^{(\mathrm{B})}$.) Several quantitative measures of entanglement can be extracted from this eigenvalue spectrum[101,242]. These are usually the different kinds of

entropies of the reduced density matrix, characterizing its mixedness. The most commonly used measure is the *von Neumann entropy*[179,189]

$$S^{(\mathrm{A})} \equiv S(\rho^{(\mathrm{A})}) = -\mathrm{Tr}\rho^{(\mathrm{A})} \ln \rho^{(\mathrm{A})}, \tag{160}$$

others include the more general one-parameter family of *Rényi entropies*[101,102,242] for parameter lower than 1, the *Hartley entropy* $\ln r_{\mathrm{Sch}}$ (which can be considered as the Rényi entropy in the limit when its parameter tends to 0), the *Schmidt rank* r_{Sch} itself, the one-parameter family of *Tsallis entropies*[73], the *concurrence-squared*, or *linear entropy* (the latter two are, up to normalization, the Tsallis entropy for parameter 2). On the other hand, the von Neumann entropy is the Rényi or Tsallis entropy in the limit when their parameters tend to 1.

The definitive property, based on which the entropies are proper measures of entanglement, is the *monotonity under LOCC:* entanglement is *quantum* correlation, so any measure of entanglement must not increase under applying *Local Operations* (that is, inside subsystems) and using *Classical Communication* between subsystems[23,46,101,227,242]. Here we have to give an important remark. This locality concept is understood with respect to the notion of *subsystems*. The subsystems have very different meanings in the first- and second-quantized description of quantum systems. In the first quantized case, the subsystems are the electrons (they can occupy different orbitals), their entanglement (*particle-entanglement*) can not increase if we apply LOCC for them, for example, if we change the local basis φ_i in $\mathcal{V}^d$ from which the Slater determinants are built up (see section 2.2), especially, changing from atomic orbitals to molecular orbitals or reverse. In the second quantized case, the subsystems are the orbitals or sites (they can be occupied by electrons), their entanglement (*orbital-entanglement* or *site-entanglement*) can not increase if we apply LOCC for them (see section 2.4), for example, if we change the local basis $|\phi_{\mu_i}^{\{i\}}\rangle$ in Λ_i for a local subspace approximation. However, since the isomorphism ι in Eq. (32) is *nonlocal*, i.e., it does not respect the tensor product structure either in $\mathcal{F}^d$ or in $\Lambda^{(d)}$, a basis change in $\mathcal{V}^d$ altough does not change the particle-entanglement but does change the orbital-entanglement. (C.f. Eq. (46).) This will be utilized in section 4.4.6 for reducing the overall orbital-entanglement by changing locally the particle basis.

Once the eigenvalues $\omega_m^{(\mathrm{A})}$ of $\rho^{(\mathrm{A})}$ are known, the von Neumann entropy (160) can be calculated, leading to

$$S^{(\mathrm{A})} = -\sum_m \omega_m^{(\mathrm{A})} \ln \omega_m^{(\mathrm{A})}. \tag{161}$$

In the examples (155) above, one can conclude that the entanglement measured by the von Neumann entropy for $|\Psi_{\mathrm{ent}}\rangle$ is $S^{(\mathrm{A})}(\Psi_{\mathrm{ent}}) = \ln 2$, while for $|\Psi_{\mathrm{sep}}\rangle$ is $S^{(\mathrm{A})}(\Psi_{\mathrm{sep}}) = 0$. It turns out also that $|\Psi_{\mathrm{ent}}\rangle$ is maximally entangled in the two-electron subspace of a two-orbital system. (The base of the logarithm in the above expressions are often set to 2, in which case the von Neumann entropy is measured in the units called *qubit*, the quanum analogy of the *bit* in classical information theory.)

In Eq. (160), subsystem (A) can be formed, in general, from an arbitrary subset of the total set of orbitals. If it is only one orbital, (A) = $\{i\}$, then its entropy is called *orbital entropy*, S_i. The number of orbitals included in (A) can be tailored to obtain specific information on the distribution of entanglement, which can then be used to characterize the physical nature of the system.

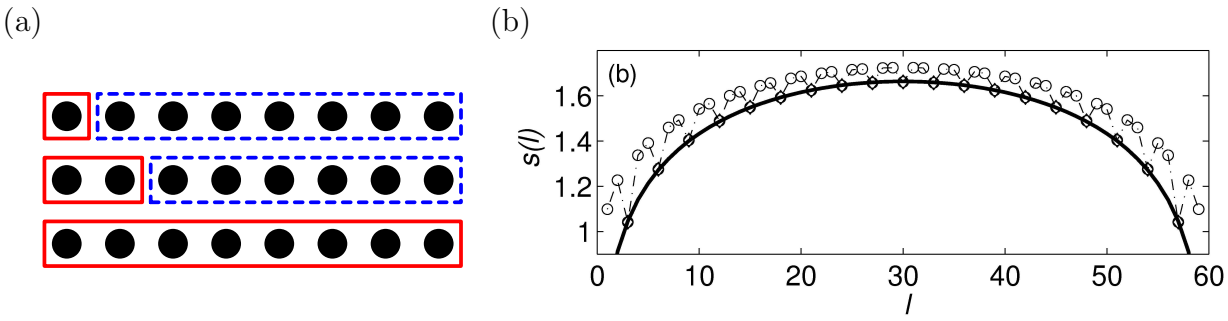


Figure 14: (a) Contiguous block of orbitals to determine block entropy. (b) Block entropy profile $S^{\{1,2,\ldots,l\}}$ obtained with the DMRG method for a one-dimensional critical model with soft modes at $k = \pm 2\pi/3$.

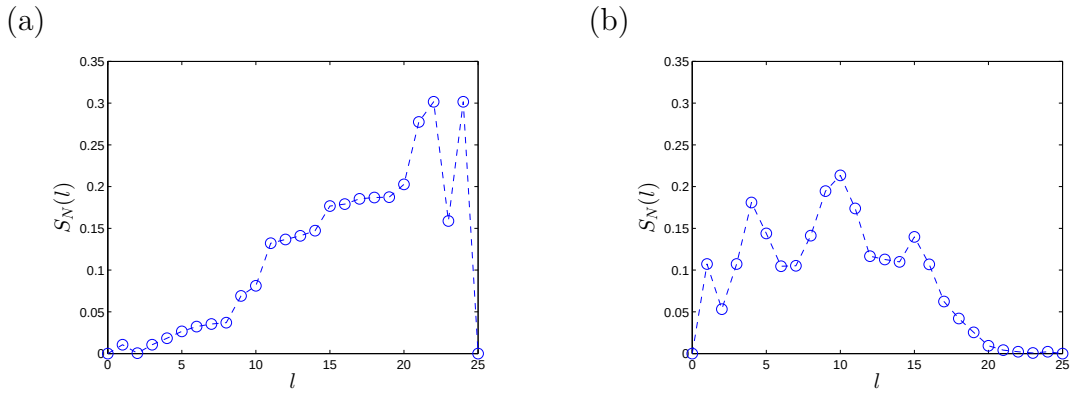


Figure 15: Block entropy profile obtained by the DMRG method for the LiF molecule at bond length $d_{\text{Li-F}} = 3.05$ a.u. for a non-optimized tensor topology (a) and for an optimized tensor topology (b).

4.2.2 Block entropy

The usual practice is to take one, two, or more neighboring orbitals into a subsystem (called also block), as is shown in Fig. 14(a) for a one-dimensional topology used in DMRG. The scaling behavior of the von Neumann entropy $S^{\{1,2,\ldots,l\}}$ of a contiguous block of the first l orbitals with the number of orbitals has also been used to study the quantum phases of one-dimensional systems. For systems with local interactions, this "block entropy" diverges logarithmically with block size l for critical systems, but saturates for gapped systems[32,241], and in certain cases its profiles provide further information about the energy spectrum[116,131]. For example, the oscillation with a period of three as is shown in Fig. 14 identifies soft modes with a wavevector, $k = \pm 2\pi/3$. In contrast to this, the block entropy has more complex behavior when non-local interactions are present[15,127] and its profile depends strongly on the ordering of the orbitals along the one dimensional chain as will be discussed below. As an example, block entropy profiles obtained with the DMRG method for the LiF molecule at bond length $d_{\text{Li-F}} = 3.05$ a.u. are shown in Fig. 15. At this point it is worth to note that not only the profiles are different but the maximum of the block entropy is much smaller in the latter case. This property will be used to optimize tensor methods as will be discussed

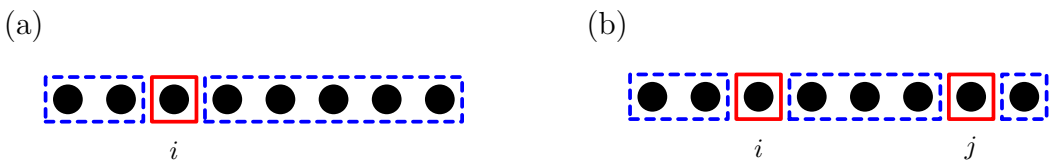


Figure 16: Partitioning of the system into single-orbital (A) = $\{i\}$ and double-orbital (A) = $\{i, j\}$ subsystems, in order to determine single-orbital entropy S_i (a) and two-orbital entropy S_{ij} (b).

below.

4.2.3 One- and two-orbital entropy and mutual information

Orbitals lying closer to and further away from the Fermi surface possess larger and smaller orbital entropy, respectively[127]. The orbital entropy is related to the mixedness of a local state and it is expressed by the eigenvalues of the one-orbital reduced density matrix (as shown in (161)) for a given orbital (A) = $\{i\}$, as shown in Fig. 16(a). Namely,

$$S_i = -\sum_{\alpha} \omega_{\alpha,i} \ln \omega_{\alpha,i}, \tag{162}$$

where $i = 1, \ldots, d$ is the orbital index, while $\omega_{\alpha,i}$ for $\alpha = 1, \ldots, q$ stands for the eigenvalues of the reduced density matrix of orbital i. The amount of contribution to the total correlation energy of an orbital can be detected by the single-orbital entropy. Since the total system is in a pure state, i.e., we calculate the ground state or an excited state, the sum of all single-orbital entropy,

$$I_{\text{tot}} = \sum_{i} S_i, \tag{163}$$

gives the amount of *total correlation* encoded in the wavefunction[129,133]. Since the full system is in a pure state, this is equal to the *total entanglement* encoded in the state/wavefunction. This quantity can be used to monitor chagnes in entanglement as system parameters are adjusted, for example, changing bond length or other geometrical properties[62,70,170].

A useful quantity to numerically characterize all kinds of correlations between pairs of orbitals is the *mutual information*

$$I_{ij} = S_i + S_j - S_{ij}, \tag{164}$$

calculated between two generally placed orbitals, i and j as shown in Fig. 16(b). Here S_i is the von Neumann entropy, Eq.(160), for a subsytem (A) chosen to be the single orbital i, and S_{ij} is the entropy for (A) chosen to consist of orbitals i and j. The mutual information I_{ij} describes the correlation between the two selected subsystems, orbitals i and j, embedded in a larger system. I_{ij} yields a weighted graph of the overall correlation of both classical and quantum origin among the orbitals. The mutual information defined in this way has been introduced previously to study correlation between neighboring orbitals in spin and fermionic chains with local interactions[130] and in quantum chemical problems in order to optimize the network structure[15,201] as well as to study molecular bonding properties in

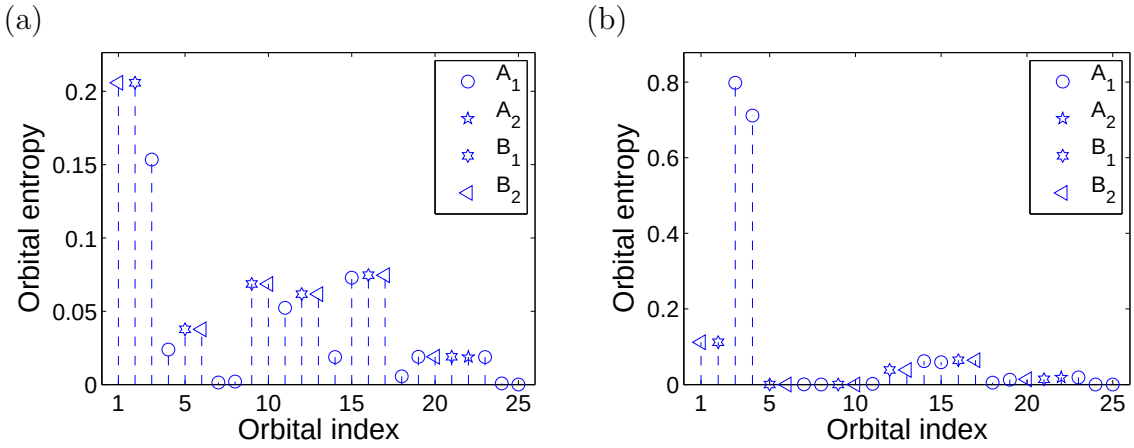


Figure 17: One orbital entropy profile for the LiF molecule at bond length (a) $_{\text{Li-F}} = 3.05$ a.u. and at (b) $d_{\text{Li-F}} = 13.7$ a.u. Symbols label the irreducible representations of the molecular orbitals in the C_{2v} point group.

various transition metal complexes[15,26,29,30]. Therefore, these quantities provide chemical information about the system, especially about bond formation and nature of static and dynamic correlation[15,29,30,62,112]. As an example, S_i and I_{ij} are shown in Figs. 17 and 18, respectively, for the equilibrium bond length $d_{\text{Li-F}} = 3.05$ a.u. and at large separation $d_{\text{Li-F}} = 13.7$ a.u.. It is clear form Fig. 18 that some orbitals are strongly entangled with several other orbitals while some orbitals are entangled with only a few others and some are almost disentangled from the system.

4.2.4 One- and two-orbital reduced density matrix and generalized correlation functions

It has been shown[17,70] that one can also analyze the sources of entanglement encoded in I_{ij} by studying the behavior of the matrix elements of the two-orbital reduced density matrix ρ_{ij}. The d-orbital wave function can be written in terms of the single-orbital q-dimensional basis as

$$|\Psi\rangle = \sum_{\alpha_1,\ldots,\alpha_d} U(\alpha_1,\ldots\alpha_d)|\phi_{\alpha_1}^{\{1\}}\rangle \otimes \ldots \otimes |\phi_{\alpha_d}^{\{d\}}\rangle, \tag{165}$$

where the α_j labels single-orbital basis states and the set of coefficients $U(\alpha_1,\ldots,\alpha_d)$ is viewed as a tensor of order d. The one- and two-orbital reduced density matrices $\rho_i = \text{Tr}_{1,\ldots,\not{i},\ldots,d}|\Psi\rangle\langle\Psi|$ and $\rho_{ij} = \text{Tr}_{1,\ldots,\not{i},\ldots,\not{j},\ldots,d}|\Psi\rangle\langle\Psi|$ can be calculated by taking the appropriate

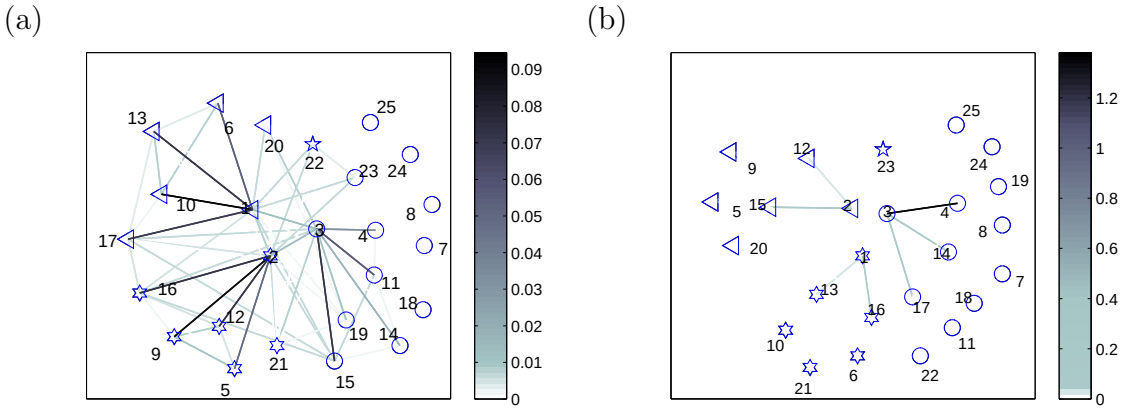


Figure 18: (Color online) Mutual information represented as a two-dimensional weighted graph for the LiF molecule at bond length (a) $d_{\text{Li-F}} = 3.05$ a.u. and at (b) $d_{\text{Li-F}} = 13.7$ a.u. Colors indicate different strengths of I_{ij} and the symbols label the irreducible representations of the molecular orbitals in the C_{2v} point group.

partial traces of $|\Psi\rangle\langle\Psi|$, leading to the matrix elements

$$\begin{aligned}\rho_i(\alpha_i, \alpha_i') &= \langle \phi_{\alpha_i}^{\{i\}} | \varrho_i | \phi_{\alpha_i'}^{\{i\}} \rangle \\ &= \sum_{\alpha_1,\ldots,\cancel{\alpha_i},\ldots,\alpha_d} U(\alpha_1,\ldots,\alpha_i,\ldots,\alpha_d)\overline{U(\alpha_1,\ldots,\alpha_i',\ldots,\alpha_d)}, \end{aligned} \tag{166a}$$

$$\begin{aligned}\rho_{ij}(\alpha_i, \alpha_j, \alpha_i', \alpha_j') &= \langle \phi_{\alpha_i}^{\{i\}} \phi_{\alpha_j}^{\{j\}} | \varrho_{ij} | \phi_{\alpha_i'}^{\{i\}} \phi_{\alpha_j'}^{\{j\}} \rangle \\ &= \sum_{\substack{\alpha_1,\ldots,\cancel{\alpha_i},\ldots,\\ \cancel{\alpha_j},\ldots,\alpha_d}} U(\alpha_1,\ldots,\alpha_i,\ldots,\alpha_j,\ldots,\alpha_d)\overline{U(\alpha_1,\ldots,\alpha_i',\ldots,\alpha_j',\ldots,\alpha_d)}. \end{aligned} \tag{166b}$$

The dimension of U grows exponentially with system size d, thus, such full tensor representations of the wave function, needed for the computation of the reduced density matrices above, are only possible for small system sizes. Using the methods described in the previous and following sections, the dth-order tensor U can, in many cases, be efficiently factorized into a product of matrices, as e.g., in (152g)

$$U(\alpha_1,\ldots,\alpha_d) = \mathbf{A}_1(\alpha_1)\mathbf{A}_2(\alpha_2)\ldots\mathbf{A}_d(\alpha_d), \tag{167}$$

leading to an MPS representation of the wave function, where the $\mathbf{A}_i(\alpha_i)$ are $M \times M$ matrices in general[238]. For systems with open boundary conditions, $\mathbf{A}_1(\alpha_1)$ and $\mathbf{A}_d(\alpha_d)$ are row and column vectors, respectively. In the MPS representation, the calculation of ρ_i and ρ_{ij} by means of Eqs. (166a) and (166b) corresponds to the contraction of the network over all states except those at orbital i in the first case and at orbital i and j in the second, as depicted in Fig. 19 for a chain with $d = 8$ orbitals.

From a different point of view, the matrix elements of ϱ_i and ϱ_{ij} in Eqs. (166a) and (166b) can be written as expectation values of projection-like operators acting on the corresponding orbitals Let the *transition operators* be defined as

$$\mathcal{T}^{(m)} = |\phi_{\alpha'}\rangle\langle\phi_\alpha|, \qquad \text{for } m = 1,\ldots,q^2, \tag{168}$$

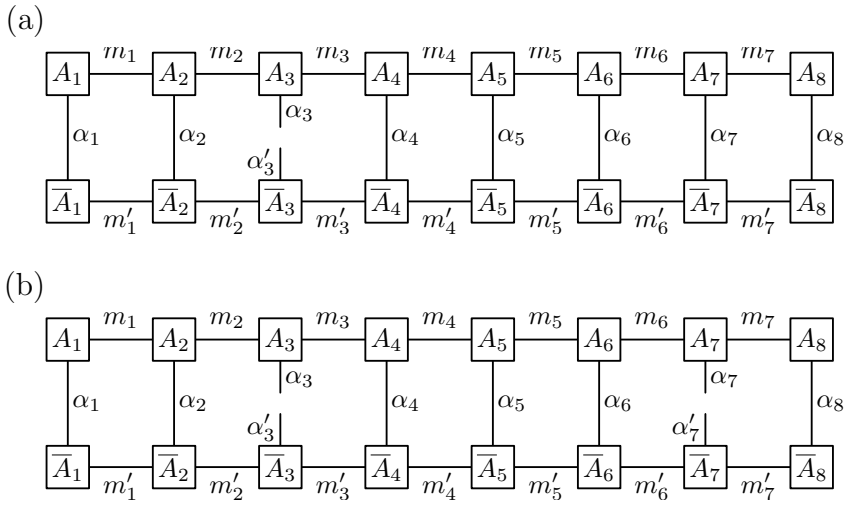


Figure 19: Contraction of the MPS network to calculate the one- (a) and two-orbital (b) reduced density matrices ϱ_i and ϱ_{ij} for a chain with $d = 8$.

which describe a possible *transition* between the initial states $|\phi_\alpha\rangle$ and the final states $|\phi_{\alpha'}\rangle$ understood for a given orbital, with the *numbering rules*

$$\alpha - 1 = ((m-1) \mod q), \qquad \alpha' - 1 = \lfloor (m-1)/q \rfloor, \tag{169a}$$

$$m - 1 = (\alpha - 1)q + \alpha' - 1. \tag{169b}$$

(Here $\lfloor x \rfloor$ denotes the floor function, the integral part of x.) These operators can be extended to operate on the complete Hilbert space consisting of d local Hilbert spaces labeled by $i = 1, \ldots, d$ as

$$\mathcal{T}_i^{(m)} = \mathbb{I} \otimes \cdots \otimes \mathbb{I} \otimes \mathcal{T}^{(m)} \otimes \mathbb{I} \otimes \cdots \otimes \mathbb{I}, \tag{170}$$

with the operator $\mathcal{T}^{(m)}$ in the i-th position.

One can now easily check that the matrix elements of the one- and two-orbital reduced density matrices, given in (166a)-(166b), can be expressed as the expectation values of the transition operators for one and for two sites, respectively, as follows

$$\rho_i(\alpha_i, \alpha_i') = \langle \mathcal{T}_i^{(m_i)} \rangle, \tag{171a}$$

$$\rho_{ij}(\alpha_i, \alpha_j, \alpha_i', \alpha_j') = \langle \mathcal{T}_i^{(m_i)} \mathcal{T}_j^{(m_j)} \rangle, \tag{171b}$$

using the numbering rules (169) for each orbitals. That is, the matrix representation of the one-orbital reduced density operator ϱ_i can be constructed from expectation values of operators describing *transitions between the single-orbital basis* $|\phi_{\alpha_i}^{\{i\}}\rangle$, while the two-orbital reduced density operator ϱ_{ij} can be constructed from expectation values of operators describing *transitions between two-orbital basis states* $|\phi_{\alpha_i}^{\{i\}} \phi_{\alpha_j}^{\{j\}}\rangle \equiv |\phi_{\alpha_i}^{\{i\}}\rangle \otimes |\phi_{\alpha_j}^{\{j\}}\rangle$. This is a generalization of the procedure introduced in the DMRG context for spin-1/2 fermion models[29,201]. In the following, we refer to the expectation values of pairs of state-transition operators in Eq. (171b) as *generalized* correlation functions in order to distinguish them from conventional correlation functions, i.e., those based on physically motivated self-adjoint operators such as local spin or density operators. For (171a), note that when the individual local basis

states are completely distinguished by abelian quantum numbers, the one-orbital density matrix is diagonal and has the form $\rho_i(\alpha_i,\alpha_i') = \delta_{\alpha_i,\alpha_i'}\langle \mathcal{T}_i^{(\alpha_i(q-1)+\alpha_i')}\rangle$, providing the spectrum immediately.

A given generalized correlation function measures the expectation value of the resonance amplitude between the initial and final states within a particular environment. In general, $\langle \mathcal{T}_i^{(m_i)}\,\mathcal{T}_j^{(m_j)}\rangle$ contains both connected and disconnected contributions between subsystems i and j. Therefore, it can, in general, scale to a finite value as the distance $l = |i-j|$ is increased, even if the physical correlation function goes to zero for large l. In order to circumvent this behavior, one generally study the connected part of the generalized correlation functions, $\langle \mathcal{T}_i^{(m_i)}\,\mathcal{T}_j^{(m_j)}\rangle_{\mathrm{C}} = \langle \mathcal{T}_i^{(m_i)}\,\mathcal{T}_j^{(m_j)}\rangle - \langle \mathcal{T}_i^{(m_i)}\rangle\langle \mathcal{T}_j^{(m_j)}\rangle$, where the disconnected part, given by the product of the expectation values of the local transition operators, is subtracted out. Note that the mutual information (164) is formulated in such a way that the disconnected parts of the generalized correlation functions do not contribute. These can be used to identify the relevant physical processes that lead to the generation of the entanglement[17,70].

As an example, let us take the spin-1/2 fermionic model. Here the single-electron basis states can be empty, occupied with a single spin-down or spin-up electron, or doubly occupied, with the corresponding basis states denoted as $|-\rangle$, $|\downarrow\rangle$, $|\uparrow\rangle$, and $|\uparrow\downarrow\rangle$, as before. Since the local basis is $q = 4$-dimensional, $q^2 = 16$ possible transition operators $\mathcal{T}^{(m)}$ arise, as is displayed in Table. 4. They can be written explicitly in terms of local fermion creation $\mathbf{c}_{i,s}^\dagger$, annihilation $\mathbf{c}_{i,s}$ and number $\mathbf{n}_{i,s}$ operators (51c)-(51d) as

$$
\begin{aligned}
\mathcal{T}^{(1)} &= (\mathbb{I}-\mathbf{n}_\uparrow)(\mathbb{I}-\mathbf{n}_\downarrow), & \mathcal{T}^{(2)} &= (\mathbb{I}-\mathbf{n}_\uparrow)\mathbf{c}_\downarrow,\\
\mathcal{T}^{(3)} &= \mathbf{c}_\uparrow(\mathbb{I}-\mathbf{n}_\downarrow), & \mathcal{T}^{(4)} &= -\mathbf{c}_\uparrow\mathbf{c}_\downarrow,\\
\mathcal{T}^{(5)} &= (\mathbb{I}-\mathbf{n}_\uparrow)\mathbf{c}_\downarrow^\dagger, & \mathcal{T}^{(6)} &= (\mathbb{I}-\mathbf{n}_\uparrow)\mathbf{n}_\downarrow,\\
\mathcal{T}^{(7)} &= -\mathbf{c}_\uparrow\mathbf{c}_\downarrow^\dagger, & \mathcal{T}^{(8)} &= \mathbf{c}_\uparrow\mathbf{n}_\downarrow,\\
\mathcal{T}^{(9)} &= \mathbf{c}_\uparrow^\dagger(\mathbb{I}-\mathbf{n}_\downarrow), & \mathcal{T}^{(10)} &= \mathbf{c}_\uparrow^\dagger\mathbf{c}_\downarrow,\\
\mathcal{T}^{(11)} &= \mathbf{n}_\uparrow(\mathbb{I}-\mathbf{n}_\downarrow), & \mathcal{T}^{(12)} &= -\mathbf{n}_\uparrow\mathbf{c}_\downarrow,\\
\mathcal{T}^{(13)} &= \mathbf{c}_\uparrow^\dagger\mathbf{c}_\downarrow^\dagger, & \mathcal{T}^{(14)} &= \mathbf{c}_\uparrow^\dagger\mathbf{n}_\downarrow,\\
\mathcal{T}^{(15)} &= -\mathbf{n}_\uparrow\mathbf{c}_\downarrow^\dagger, & \mathcal{T}^{(16)} &= \mathbf{n}_\uparrow\mathbf{n}_\downarrow.
\end{aligned}
\tag{172}
$$

The non-vanishing matrix elements of the two-orbital density matrix ρ_{ij} are given in Table 5. Note that the two-orbital density matrix is block-diagonal in the particle number N_c and in the z component of the spin S^z. The block-diagonal structure is evident, and the values of m_i and m_j appropriate for each matrix element are displayed.

Illustrating these, some generalized correlation functions are plotted for the LiF molecule in Fig. 20. As was mentioned in the begining of this section[17,70], the generalized correlation functions (matrix elements for ϱ_{ij}) are connected to the values of the mutual information I_{ij}, which is plotted in Fig. 34 later.

	$\|-\rangle_i$	$\|\downarrow\rangle_i$	$\|\uparrow\rangle_i$	$\|\uparrow\downarrow\rangle_i$
$\|-\rangle_i$	$\mathcal{T}_i^{(1)}$	$\mathcal{T}_i^{(2)}$	$\mathcal{T}_i^{(3)}$	$\mathcal{T}_i^{(4)}$
$\|\downarrow\rangle_i$	$\mathcal{T}_i^{(5)}$	$\mathcal{T}_i^{(6)}$	$\mathcal{T}_i^{(7)}$	$\mathcal{T}_i^{(8)}$
$\|\uparrow\rangle_i$	$\mathcal{T}_i^{(9)}$	$\mathcal{T}_i^{(10)}$	$\mathcal{T}_i^{(11)}$	$\mathcal{T}_i^{(12)}$
$\|\uparrow\downarrow\rangle_i$	$\mathcal{T}_i^{(13)}$	$\mathcal{T}_i^{(14)}$	$\mathcal{T}_i^{(15)}$	$\mathcal{T}_i^{(16)}$

Table 4: Single-orbital operators describing transitions between single-orbital basis states for a $S = 1/2$ spin system.

$\rho_{i,j}$	n=0, s_z=0	n=1, s_z=-$\frac{1}{2}$		n=1, s_z=$\frac{1}{2}$		n=2, s_z=-1	n=2, s_z=0				n=2, s_z=1	n=3, s_z=-$\frac{1}{2}$		n=3, s_z=$\frac{1}{2}$		n=4, s_z=0
	$--$	$-\downarrow$	$\downarrow-$	$-\uparrow$	$\uparrow-$	$\downarrow\downarrow$	$-\uparrow\downarrow$	$\downarrow\uparrow$	$\uparrow\downarrow$	$\uparrow\downarrow-$	$\uparrow\uparrow$	$\downarrow\uparrow\downarrow$	$\uparrow\downarrow\downarrow$	$\uparrow\uparrow\downarrow$	$\uparrow\downarrow\uparrow$	$\uparrow\downarrow\uparrow\downarrow$
$--$	(1,1)															
$-\downarrow$		(1,6)	(2,5)													
$\downarrow-$		(5,2)	(6,1)													
$-\uparrow$				(1,11)	(3,9)											
$\uparrow-$				(9,3)	(1,11)											
$\downarrow\downarrow$						(6,6)										
$-\uparrow\downarrow$							(1,16)	(2,15)	(3,14)	(4,13)						
$\downarrow\uparrow$							(5,12)	(6,11)	(7,10)	(8,9)						
$\uparrow\downarrow$							(9,8)	(10,7)	(11,6)	(12,5)						
$\uparrow\downarrow-$							(13,4)	(14,3)	(15,2)	(16,1)						
$\uparrow\uparrow$											(11,11)					
$\downarrow\uparrow\downarrow$												(6,16)	(8,14)			
$\uparrow\downarrow\downarrow$												(14,8)	(16,6)			
$\uparrow\uparrow\downarrow$														(11,16)	(12,15)	
$\uparrow\downarrow\uparrow$														(15,12)	(16,11)	
$\uparrow\downarrow\uparrow\downarrow$																(16,16)

Table 5: The two-orbital reduced density matrix ρ_{ij} for SU(2) fermions expressed in terms of single-orbital operators, $\mathcal{T}_i^{(m_i)}$ with $m_i = 1, \ldots, 16$. For better readability only the operator number indices m are shown, that is, (m_i, m_j) corresponds to $\langle \mathcal{T}_i^{(m_i)} \mathcal{T}_j^{(m_j)} \rangle$. Here N_c and S^z denote the particle-number and z spin component quantum numbers of the two orbitals.

4.3 Methods based on block transfromation procedures

4.3.1 Block renormalization group method (BRG)

One of the first attempts to approximate the full configuration Hilbert space $\Lambda^{(d)} = \otimes_{i=1}^{d} \Lambda_i$ ($\dim \Lambda_i = q$) of a d-orbital system goes back to the late 1960's when Kadanoff invented the *Block Spin Renormalization Group* method and applied it to the two-dimensional Ising model[107]. This was later extended to quantum systems in one dimension called *Block Renormalization Group* (BRG) method[59,106]. The main idea of the method is to group d_s number of orbitals into blocks. The total Hamiltonian is then written as a sum of terms corresponding to the interactions within the blocks (*intrablock* Hamiltonian) and terms corresponding to the interactions between the blocks (*interblock* Hamiltonian). The unitary matrix $\mathbf{O}$ introduced in Sec. 4.1.3 is formed from the q lowest eigenstates of the intrablock Hamiltonian and operators are transformed to a new basis using Eq. 147. Using the transformed operators the interblock Hamiltonian can also be expressed. Truncating the Hilbert space of the blocks and keeping only q states per block ensures that one can rescale the interaction strengths (*flow equations*) and thus the original form of the Hamiltonian is retained. In the next iteration step the d_s-blocks are collected. The schematic plot of the procedure is shown in Fig 21(a). The procedure is repeated until subsequent iterations do not change the

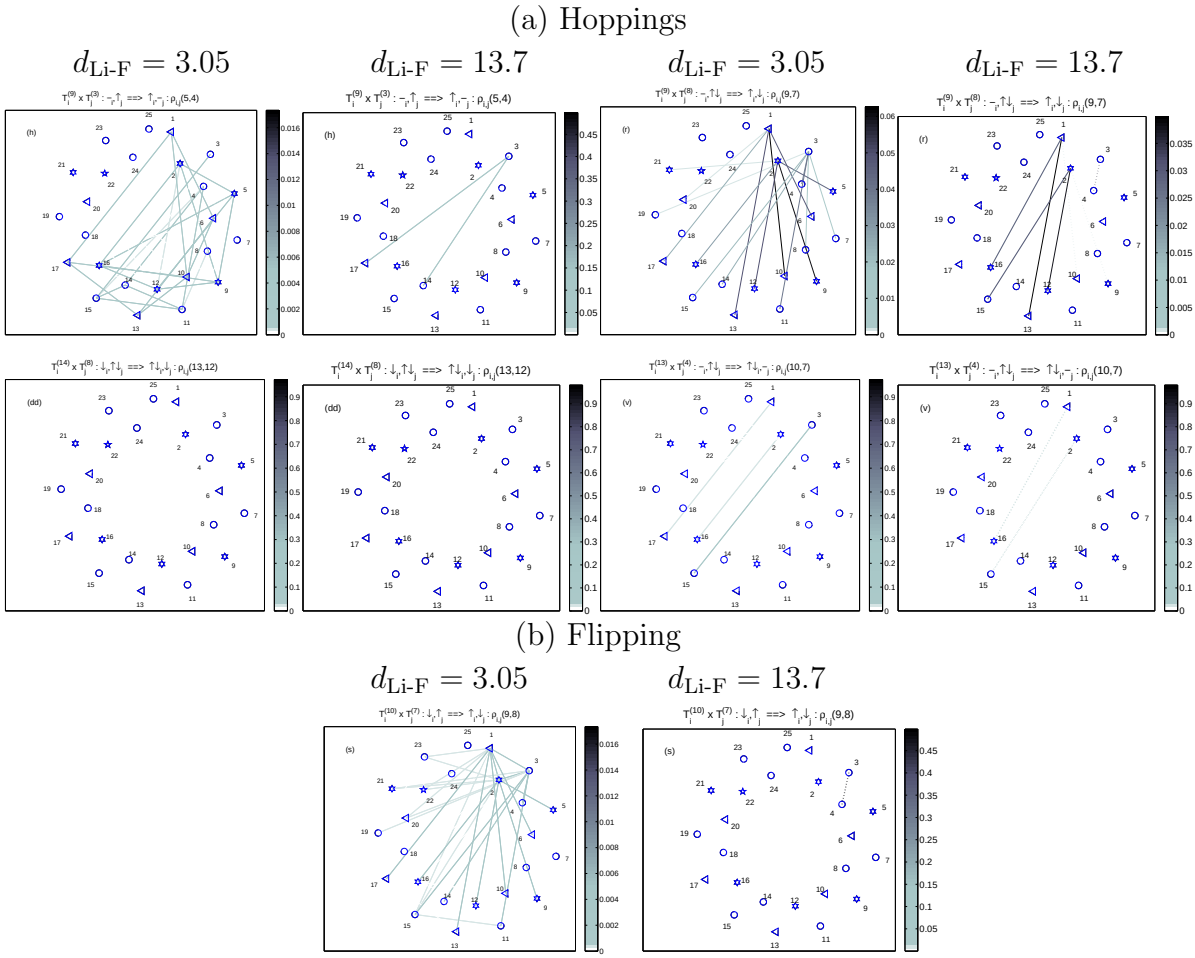


Figure 20: Pictorial representation of the absolute value of the generalized correlation functions used to construct the lower-triangular elements of the two-orbital reduced density matrix for LiF at $d_{\text{Li-F}} = 3.05$ a.u. and at $d_{\text{Li-F}} = 13.7$ a.u.. Strength of transition amplitues between initial ($|\phi^{\{i\}}_{\alpha_i}\phi^{\{j\}}_{\alpha_j}\rangle$) and final states ($|\phi^{\{i\}}_{\alpha'_i}\phi^{\{j\}}_{\alpha'_j}\rangle$) on orbital i and j are indicated with different line colors. Note the different scales used for colorbars in case of the various figures.

interaction strengths, i.e., until the so-called *fixed point* of the RG transformation is reached when measurable quantities corresponding to the $d \to \infty$ limit can be calculated. While this method gave reasonably good results for some one-dimensional models with local interactions, using such systematic change of basis and truncation led to loss of information in each iteration step and the accumulation of the error hindered the application of the method for more complex problems. In case of systems with finite number of orbitals this block transformation procedure can also be carried out until all orbitals are included in a single block and the approximated ground state energy can be calculated. This corresponds to the root for the Hierarchial Tucker format discussed in Sec. 3.6. Due to the dramatic truncation of the states and non-local interactions this procedure cannot be applied efficiently in quantum chemistry. However, the BRG method also serves the basis of hierarchal tensor representation and tree tensor network state ansatz discussed in Sec. 4.3.4. Recently, extension of the method known as the *Multiscale Entanglement Renormalization Ansatz*[67] (MERA) gave a

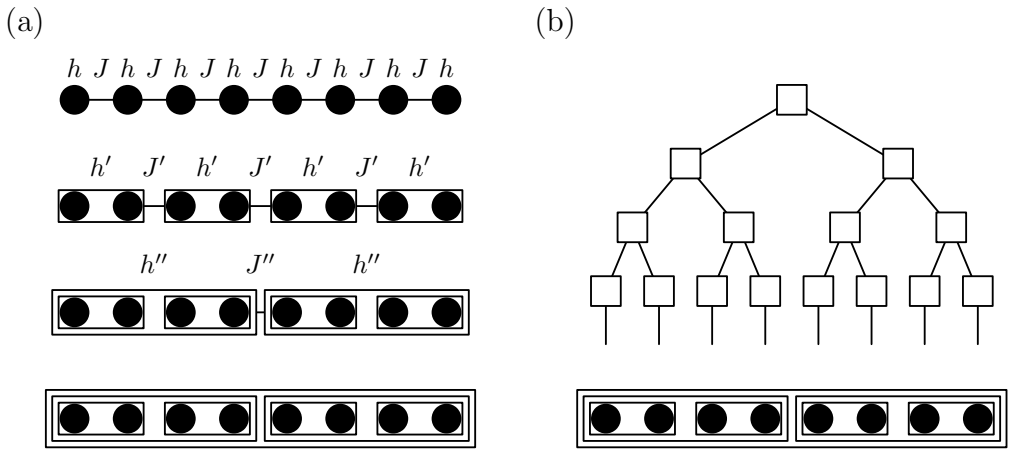


Figure 21: Schematic plot of the Block Renormalization Group (BRG) method as bloc transfromation procedure where h and J label on-orbital and nearest neighbor interaction, respectively (a), and as a tree-network (b).

new impetus to its application for strongly correlated systems.

4.3.2 Numerical renormalization group method (NRG)

Another variant of the RG method, known as the *Numerical Renormalization Group* (NRG) method shown in Fig. 22 is due to Wilson[258]. In the NRG related Hamiltonian an impurity interacts with a local fermion. The dynamics of this fermion is described by a semi-infinite one dimensional network, also know as the Wilson chain. The impurity sits on the left side and electrons can move along the chain with an exponentially decreasing hopping amplitude $\lambda^{-j/2}$. Therefore, each orbital represents a different energy scale. Starting with the very left orbital, new blocks including l orbitals are formed by adding orbitals systematically to the block, i.e., $\Xi^{(\mathrm{L})} = \Xi^{(\mathrm{l})} \otimes \Lambda_{l+1}$ where in the first step $\Xi^{(\mathrm{l})} = \Lambda_1$. In each iteration step the block Hamiltonian is solved and the unitary transformation matrix $\mathbf{O}$ is formed from eigenstates corresponding to the lowest M eigenvalues. The block Hamiltonian is rescaled based on the decay rate of the hopping and the intrablock Hamiltonian is determined on the new basis. Another major difference compared to the BRG method is that in NRG $q < M \ll q^d$ states are kept, thus the original form of the Hamilton is lost. Due to the appearance of new operators during the iteration scheme flow equations described above cannot be studied. The change in the energy spectrum, however, can be analyzed and once subsequent iterations leave the spectrum unchanged the fix point is reached. This approach works well due to the separation of energy scales. A problem, however, arises for lattice models when $\lambda \to 1$ and error starts to accumulate significantly for increasing block size. This hindered the application of NRG to large lattice models. Quite recently, an extension of the method using a similar blocking structure as in DMRG has led to the development of the so called density matrix numerical renormalization group (DM-NRG) which allows us to study more complex problems[12,98,188,231,246].

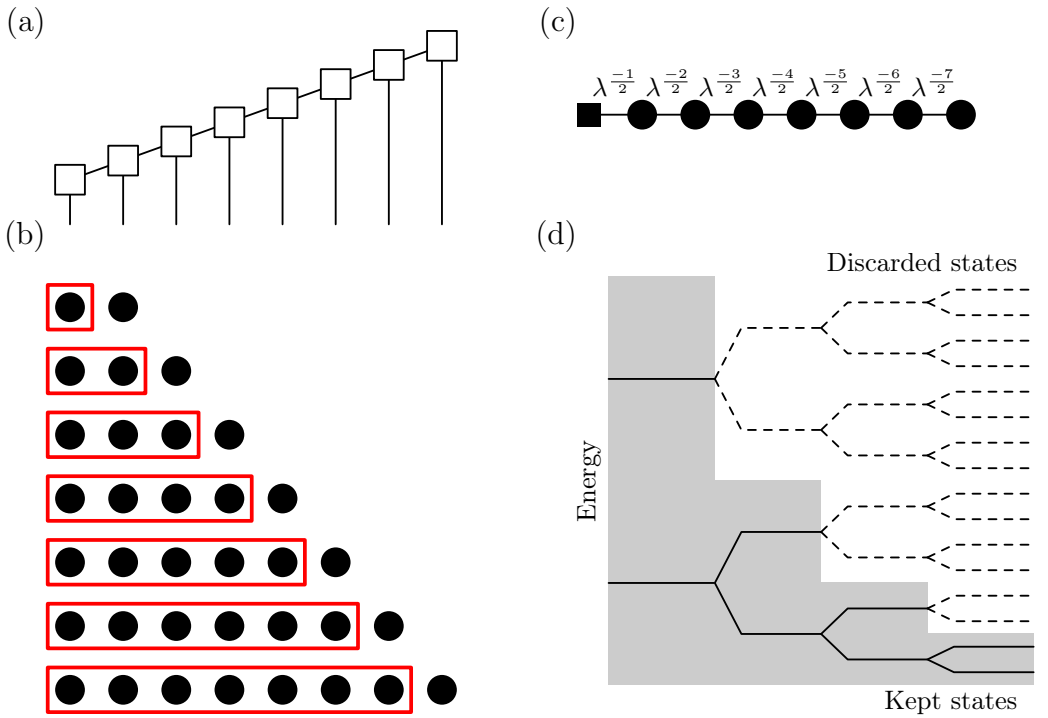


Figure 22: Schematic plot of the Numerical Renormalization Group (NRG) method a block-decimization procedure (b) leading to a tree-network (a). Hamiltonian on the Wilson chain of length d: the hopping is decreasing exponentially (c). A complete basis of a Wilson chain represented as the exponentially increasing number of energy levels belonging to the successive iterations. Continuous/dashed lines represent kept, low-energy/discarded, high-energy levels, respectively. For the consecutive iteration steps the distances between the levels illustrates how the energy resolution of NRG gets exponentially refined (d).

4.3.3 Density matrix renormalization group method (DMRG)

In order to circumvent problems discussed for BRG and NRG, in the two-site variant of the *Density Matrix Renormalization Group* (DMRG) method[249] $\Lambda^{(d)}$ is approximated by a tensor product space of four tensor spaces, i.e., $\Xi^{(d)}_{\mathrm{DMRG}} = \Xi^{(\mathrm{l})} \otimes \Lambda_{l+1} \otimes \Lambda_{l+2} \otimes \Xi^{(\mathrm{r})}$. This is called superblock and the basis states of the blocks are optimized by successive application of the singular value decomposition as discussed in Secs. 3.5 and 4.2.1. Here we use the convenient notations that the whole system, consisting of d orbitals $1, 2, \ldots d$, is partitioned into blocks (subsystems), for which we use the labels (L), (l), (R) and (r). (l) simply means the block composed of the first l orbitals, that is, (l) $= \{1, 2, \ldots, l\}$. An extended block composed of the first $l+1$ orbitals is denoted as (L) $= \{1, 2, \ldots, l, l+1\}$. The other part of the system is (R) $= \{l+2, l+3, \ldots, d\}$, while (r) $= \{l+3, \ldots, d\}$. The d-orbital wavefunction is, therefore, written as

$$|\Psi_{\mathrm{DMRG}}\rangle = \sum_{m_{(\mathrm{l})}\alpha_{l+1}\alpha_{l+2}m_{(\mathrm{r})}} U_{\mathrm{DMRG}}(m_{(\mathrm{l})}, \alpha_{l+1}, \alpha_{l+2}, m_{(\mathrm{r})}) |\xi^{(\mathrm{l})}_{m_{(\mathrm{l})}}\rangle \otimes |\phi^{\{l+1\}}_{\alpha_{l+1}}\rangle \otimes |\phi^{\{l+2\}}_{\alpha_{l+2}}\rangle \otimes |\xi^{(\mathrm{r})}_{m_{(\mathrm{r})}}\rangle \tag{173}$$

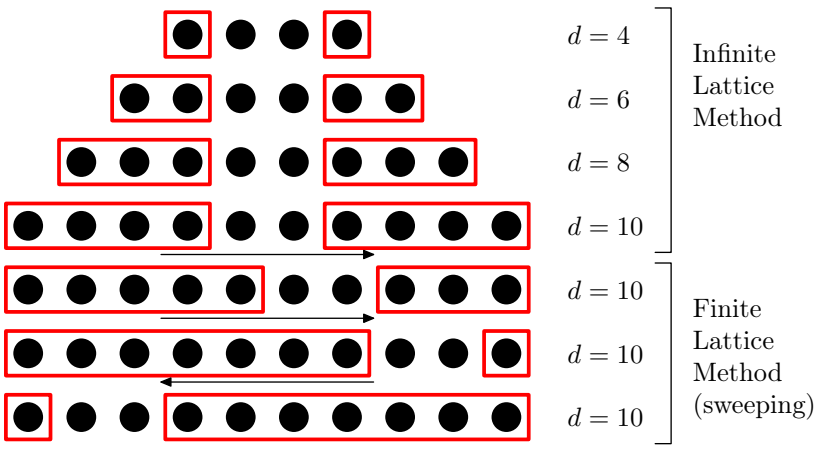


Figure 23: Decomposition of the d-orbital Hilbert space into four subsystems called superblock. The d-orbital Hilbert space is built iteratively from a left block including l active orbitals and the right block from r active orbitals. The size of the two blocks is increased in each iteration step until $l + 2 + r = d$. In the following steps the d-orbital system is partitioned asymmetrically, i.e. the size of left block is increased systematically while the size of the right block is decreased until $l = d-3$ and $r = 1$. The same procedure is repeated in the opposite direction until $l = 1$ and $r = d - 3$. This procedure is called sweeping (macro-iteration step).

where the tensor U_{DMRG} is determined by an iterative diagonalization of the corresponding so called superblock Hamiltonian. The dimensions of the spaces of the local *left block* including l orbitals and the *right block* with $r = d - l - 2$ orbitals are denoted with $M_l = \dim \Lambda^{(\mathrm{l})}$ and $M_r = \dim \Lambda^{(\mathrm{r})}$, respectively. Since $\dim \Lambda_{l+1} = \dim \Lambda_{l+2} = q$, the resulting dimensionality of the DMRG wave function is $\dim \Xi^{(d)}_{\mathrm{DMRG}} = q^2 M_l M_r \ll q^d$.

In the *original version* of the DMRG, introduced to treat finite one-dimensional lattice models[249], the Hilbert space of a lattice with d sites is built iteratively starting with four sites as shown in Fig. 23. In each iteration step, the Hilbert space $\Xi^{(\mathrm{L})}$ of an enlarged block (L) is formed from the tensor product of the Hilbert spaces of the block $\Xi^{(\mathrm{l})}$ and the adjacent site Λ_{l+1} – similarly $\Xi^{(\mathrm{R})}$ from Λ_{l+2} and $\Xi^{(\mathrm{r})}$ – and transformed to a new *truncated* basis by using a unitary operation based on singular value decomposition as discussed in section 4.2.1. Therefore, in each iteration step the size of the effective system is increased by two until the desired length d is achieved. This procedure is called *infinite-lattice procedure*. In the following steps the d-site system is partitioned asymmetrically, i.e. the size of left block is increased systematically while the size of the right block is decreased until $l = d - 3$ and $r = 1$. In each iteration step, the approximated Hilbert space of the left block (called *system block*) is improved as it interacts with the right block (called *environment*). The same procedure is repeated in the opposite direction until $l = 1$ and $r = d - 3$ when the left block becomes the environment block and the right block the system block. This procedure is called *sweeping* (macro-iteration) and it is a part of the so called *finite-lattice method*. For more detailed derivations we refer to the original papers and review articles[211,249,250].

In analogy, in the infinite-lattice procedure one can say that the d-orbital Hilbert space is built iteratively by forming l-orbital and r-orbital blocks from the one-orbital Hilbert

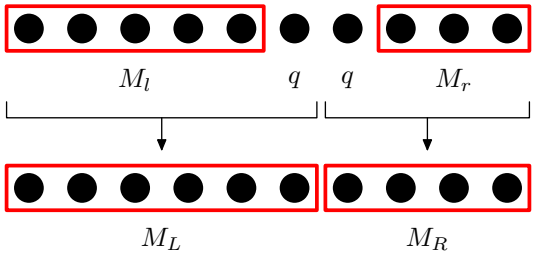


Figure 24: Schematic plot of a DMRG iteration step in order to increase block sizes and form a bipartite representation for the singular value decomposition. (l) and (r) denote the left and right block of length l and r, and of dimension M_l and M_r, respectively, $\bullet$ stands for the intermediate orbitals ($\{l+1\}$ and $\{l+2\}$) with dimension q. The blocks (L) = (l)$\bullet$, (R) = $\bullet$(r) have dimension M_L and M_R, respectively.

spaces starting with an *"active space"* including only four orbitals. In each iteration step the number of active orbitals is increased by two until all the d orbitals become active, i.e., part of either the left or right block. This procedure serves as the initialization of the MPS network with d component tensors. When the network is formed the elements of the A_i matrices are random numbers. The infinte lattice method can be viewed as a procedure to start with four "active" component tenors by setting the remaining $d-4$ component tensors to trivial. This means that the m_{i-1} and m_i indices of the corresponding $A_i(m_{i-1}, \alpha_i, m_i)$ takes only the value 1, and $A_i(1, \alpha_i, 1) = \delta_{1,\alpha_i}$ that is, $A_i(1,1,1) = 1$ and the others are 0. In each iteration step, the number of "active" component tensors is increased by two until no component tensors are set to trivial.

In quantum chemistry, it is more efficient to start with an initial network which already corresponds to the finite system with d orbitals as has been introduced through the *Dynamically Extended Active Space* (DEAS) procedure[127]. In the DEAS procedure one starts with a superblock structure with $l = 1$ and $r = d - 3$, as is shown in Fig. 25, and carries out the forward and backward sweeping procedure, i.e., the finite lattice method as described above. A cruical problem, however, is that during the first sweep when the left block is optimized the right block Hilbert space has to be approximated with $M_r \ll q^r$ basis states. An efficient method to carry out such optimization will be discussed in Sec. 4.4.7 based on the Configuration Interaction (CI) procedure.

Let us highlight the main aspect of DMRG procedure once again: If one could represent the Hilbert spaces of the four subsystems used in the two-site DMRG exactly using one-orbital basis states then in the first step of the DEAS procedure this would mean $M_l = q$ and $M_r = q^{d-3}$ and $\Lambda^{(d)}_{\mathrm{DMRG}} = \Lambda^{(\mathrm{l})} \otimes \Lambda_{l+1} \otimes \Lambda_{l+2} \otimes \Lambda^{(\mathrm{r})}$. By traversing through the system back-and-forth the left and right block Hilbert spaces are transformed and truncated, and after a full sweep the approximated subspace is given as $\Xi^{(d)}_{\mathrm{DMRG}} = \Xi^{(\mathrm{l})} \otimes \Lambda_{l+1} \otimes \Lambda_{l+2} \otimes \Xi^{(\mathrm{r})}$. Therefore, the d-orbital wavefunction written in terms of one-orbital basis is converted to an approximated multi-orbital basis in $\Xi^{(d)}_{\mathrm{DMRG}} = \Xi^{(\mathrm{l})} \otimes \Lambda_{l+1} \otimes \Lambda_{l+2} \otimes \Xi^{(\mathrm{r})}$, where $\dim \Xi^{(d)}_{\mathrm{DMRG}} \ll \dim \Lambda^{(d)}$ depending on the level of truncation.

A main difference compared to the BRG and NRG methods is how the transformation matrix $\mathbf{O}$ is constructed. In a given iteration step (see Fig. 24) the (l)$\bullet$ composite system is combined to one subsystem (L) with $\Xi^{(\mathrm{L})} = \Xi^{(\mathrm{l})} \otimes \Lambda_{l+1}$ and $\bullet$(r) to another one (R) with

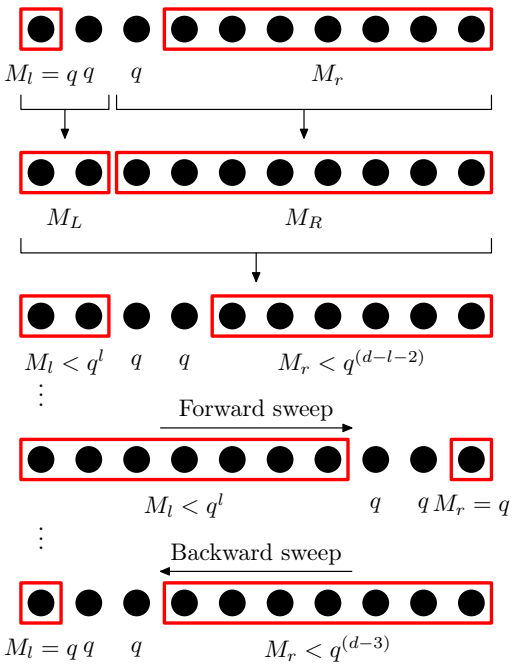


Figure 25: A modified initialization of the tensor network with d orbitals used in the Dynamically Extended Active Space (DEAS) procedure. In the DEAS procedure one starts with a superblock structure with $l = 1$ and $r = d-3$ and use an approximated Hilbert space with diemnsion $M_r \ll q^r$.

$\Xi^{(\mathrm{R})} = \Lambda_{l+2} \otimes \Xi^{(\mathrm{r})}$. This leads to $\Xi^{(\mathrm{L})} \otimes \Xi^{(\mathrm{R})} = \Xi^{(d)}_{\mathrm{DMRG}} \subseteq \Lambda^{(d)}$ and the bipartite representation of the wavefunction is formed as

$$|\Psi_{\mathrm{DMRG}}\rangle = \sum_{m_{(\mathrm{L})} m_{(\mathrm{R})}} U_{\mathrm{DMRG}}(m_{(\mathrm{L})}, m_{(\mathrm{R})}) |\phi^{(\mathrm{L})}_{m_{(\mathrm{L})}}\rangle \otimes |\phi^{(\mathrm{R})}_{m_{(\mathrm{R})}}\rangle. \tag{174}$$

According to section 4.2.1, using singular value decomposition it can be written as a single sum of tensor products. The new basis states $|\xi^{(\mathrm{L})}_{m_{(\mathrm{L})}}\rangle$ and $|\xi^{(\mathrm{R})}_{m_{(\mathrm{R})}}\rangle$ given in Eq. (154) are obtained by diagonalizing the reduced subsystem density matrices $\rho^{(\mathrm{L})}$ and $\rho^{(\mathrm{R})}$, see Eq. (156). The transformation matrix $\mathbf{O}$ introduced in section 4.1.3 is formed from eigenstates $|\xi^{(\mathrm{L})}_{m_{(\mathrm{L})}}\rangle$ (or $|\xi^{(\mathrm{R})}_{m_{(\mathrm{R})}}\rangle$) corresponding to the $M^{\mathrm{kept}}_l \leq M_l q$ (or $M^{\mathrm{kept}}_r \leq M_r q$) largest eigenvalues ω_m. Due to the truncation of basis states the so-called *truncation error* is defined as the sum of the truncated number of eigenvalues of the reduced subsystem density matrix deviates from unity, i.e.,

$$\delta\varepsilon_{\mathrm{TR}} = 1 - \sum_{m=1}^{M_{\mathrm{kept}}} \omega_m. \tag{175}$$

Operators of the enlarged blocks are transformed to this new basis as $(\mathbf{X}_i \mathbf{Y}_{l+1})^{(l+1)} = \mathbf{O}(\mathbf{X}^{(\mathrm{l})}_i \otimes \mathbf{Y}^{\{l+1\}}_{l+1})\mathbf{O}^\dagger$, where $\mathbf{X}_i$ and $\mathbf{Y}_{l+1}$ are $M_l \times M_l$ and $q \times q$ matrices, respectively. The number of block states, M_l and M_r, required to achieve sufficient convergence can be regarded as a function of the level of entanglement among the molecular orbitals. Hence

the maximum number of block states $M_{\max} = \max(M_l, M_r)$ determines the accuracy of a DMRG calculation[125,211] as will be investigated in the next section.

If the transformation matrix $\mathbf{O}$ in each iteration step is reindexed according to the procedure explained in Sec. 4.1.3 and the corresponding $\mathbf{B}$ matrices are stored within a full sweep then the DMRG wavefunction for a given superblock partitioning can be written in MPS form[184,237,238] as

$$\begin{aligned}|\Psi_{\mathrm{DMRG}}\rangle = \sum_{m_{(\mathrm{l})}\alpha_{l+1}\alpha_{l+2}m_{(\mathrm{r})}} & U_{\mathrm{DMRG}}(m_{(\mathrm{l})}, \alpha_{l+1}, \alpha_{l+2}, m_{(\mathrm{r})}) \\ & \times(\mathbf{B}_l(\alpha_l)\dots\mathbf{B}_2(\alpha_2))_{m_{(\mathrm{l})};\alpha_1}(\mathbf{B}_{l+3}(\alpha_{l+3})\dots\mathbf{B}_{d-1}(\alpha_{d-1}))_{m_{(\mathrm{r})};\alpha_d} \\ & \times|\phi^{\{1\}}_{\alpha_1}\rangle\otimes\dots\otimes|\phi^{\{l+1\}}_{\alpha_{l+1}}\rangle\otimes|\phi^{\{l+2\}}_{\alpha_{l+2}}\rangle\otimes\dots\otimes|\phi^{\{d\}}_{\alpha_d}\rangle.\end{aligned} \tag{176}$$

Therefore, DMRG can be viewed as an efficient method to generate the optimized set of $\mathbf{A}_i$ ($\mathbf{B}_i$) matrices used to construct the MPS representation of the d-orbital wavefunction. Since in this representation the d-orbital wavefunction is written as a linear combination of the tensor product of the one-orbital basis (CI coefficients), it allows one to connect the DMRG wavefunction to conventional quantum chemical techniques. For example, the CI-coefficients of the most relevant terms can be determined[27].

Concluding this section, different one-dimensional representation of tensor network state algorithms, i.e, matrix product state methods have been developped in the various communities. In this one-dimensional optimization scheme the network is built from matrices. The TT and MPS approaches are "wavefunction" oriented description of the problem while DMRG is more like an "operator" representation of the problem. In the TT and MPS the physical indices are for local q dimensional tensor spaces thus operators are $q \times q$ matrices but the $\mathbf{A}(\alpha)$ matrices must be stored. The norm is calculated by simply connecting the physical indices vertically. In contrast to this, in the DMRG description when the network is separated to a left and a right part, the operators of the left and right part are represented on a multi-orbital tensor space of dimension M_l and M_r, respectively, where both are much larger than q. Therefore the corresponding matrices of dimensions $M_l \times M_l$ and $M_r \times M_r$ must be stored during the iterative minimization procedure. In the quantum chemistry framework long range Coulomb interactions are given by the 4-th order tensor V_{ijkl} of equation (44b) thus the number of renormalized operators scales as $\mathcal{O}(d^4)$. Using, however, an efficient factorization of the interaction terms distributed among the various subsystems[266], this scaling can be reduced to $\mathcal{O}(d^2)$, see in section 4.3.5. Therefore, the required memory to store operators in a given QC-DMRG iteration step assuming $M_l = M_r \equiv M$ is $\mathcal{O}(M^2d^2)$. The computational cost of a given QC-DMRG step scales as $\mathcal{O}(M^3d^2)$ and for a full sweep $\mathcal{O}(M^3d^3)$. A main advantage of the DMRG method is, however, that in each iteration step the core tensor is optimized so orthogonalization of the left and right block states are guaranteed[211].

4.3.4 Higher dimensional network: Tree tensor network state (TTNS)

A natural extension of the MPS approach is to form an ansatz state by contracting a network of higher order tensors[45,81,135,136,147,151,167,168,180,236,237,240], as discussed in Sec. 3. A special class of such ansatz states are the *Tree Tensor Network States* (TTNS)[51,169,170,172] which

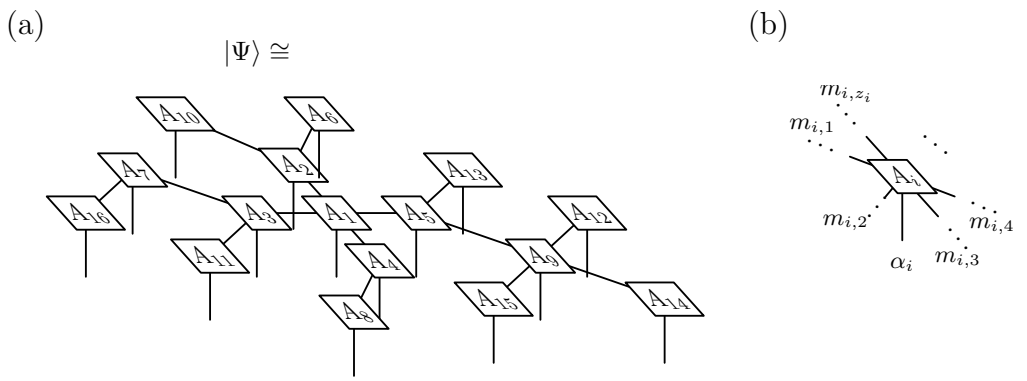


Figure 26: Schematic plot of a higher dimensional network, for example, the tree tensor network state (TTNS). Each node is represented by a tensor $\mathbf{A}_i$ of order $z_i + 1$, with z_i is an orbital dependent coordination number. The network supposed to reflect the entanglement structure of the molecule as much as possible. The vertical lines are the physical indices α_i, $i \in \{1, d\}$, while the others that connect the orbitals are virtual ones.

are formed by contracting tensors according to a tree network, as shown in Fig. 26. The structure of the tree network can be arbitrary and the coordination number can vary from site to site. Each tensor in the network represents a physical orbital and is of order $z_i + 1$, were z_i describing the coordination number of site i:

$$A_i(\alpha_i, m_{i,1}, \ldots, m_{i,z_i}). \tag{177}$$

The z_i virtual indices $m_{i,1}, \ldots, m_{i,z_i}$ are of dimension M and are contracted as the TTNS is formed. The physical index α_i is of dimension q and describes the physical state of the orbital, e.g. the number of up- and down-electrons on that orbital.

The TTNS is especially suitable to treat models in which orbitals have varying degrees of entanglement (see Figs. 17 and 18): since entanglement is transferred via the virtual bonds that connect the sites, sites with a larger coordination number are better suited to represent higher entanglement. In this way, the coordination number can be adapted according to the entanglement of the orbitals, and the orbitals can be arranged on the tree such that highly entangled orbitals are close together (see later in section 4.4.5).

An additional motivation for using a tree structure is to take advantage of the property of the tree tensor network ansatz that the long-range correlations differ from the mean-field value polynomially with distance rather than exponentially as for MPS[169]. This is due to the fact that the number of virtual bonds required to connect two arbitrary orbitals scales logarithmically with the number of orbitals d for $z > 2$, as can be seen by considering a Cayley-tree of depth Δ: the number of sites in the tree is

$$d = 1 + z \sum_{j=1}^{\Delta} (z-1)^{j-1} = \frac{z(z-1)^{\Delta} - 2}{z - 2} \tag{178}$$

and thus, the maximal distance between two orbitals, 2Δ, scales logarithmically with d for $z > 2$. On the other hand, for $z = 2$ the number of virtual bonds required to connect two arbitrary orbitals scales linearly in d.

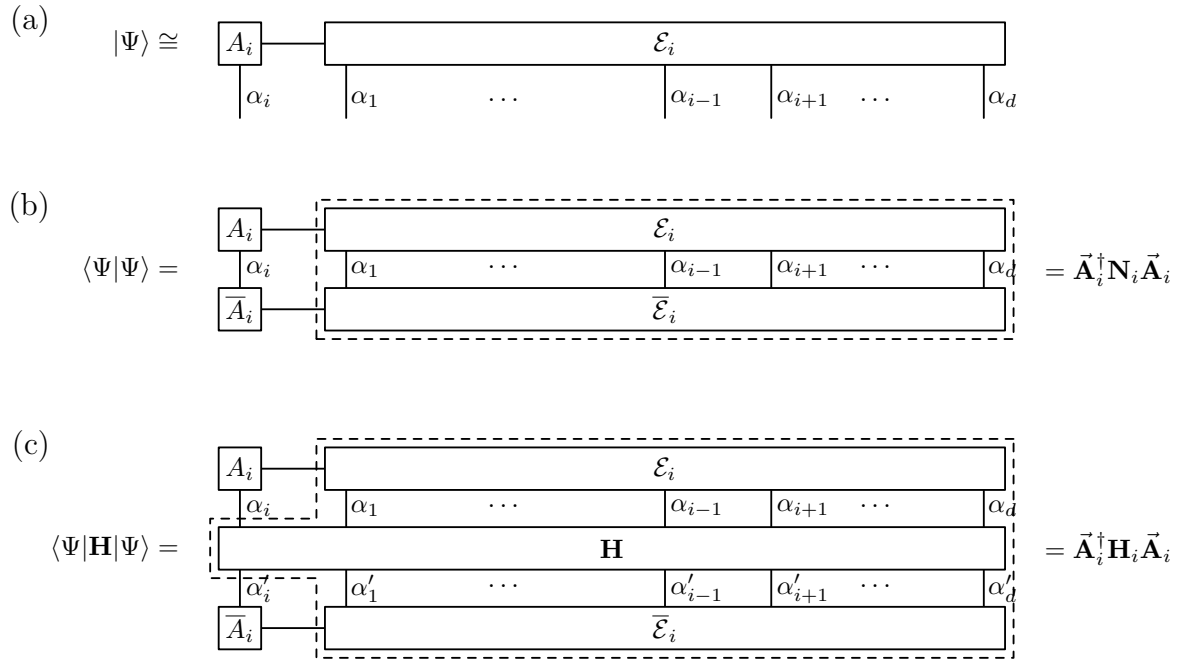


Figure 27: Concept of the variational optimization of tensor network states: (a) tensor network state $|\Psi\rangle$ separated into two parts: the tensor $\mathbf{A}_i$ that is supposed to be optimized and an environment tensor $\mathcal{E}_i$ that is formed by contracting all tensors except $\mathbf{A}_i$. (b) norm $\langle\Psi|\Psi\rangle$ of the tensor network state defined in (a); the norm equals to $\vec{\mathbf{A}}_i^\dagger \mathbf{N}_i \vec{\mathbf{A}}_i$ with $\vec{\mathbf{A}}_i$ corresponding to the qM^{z_i}-dimensional vector obtained by joining all indices of tensor $\mathbf{A}_i$, and $\mathbf{N}_i$ represents the effective environment, drawn with dashed lines. (c) expectation value $\langle\Psi|H|\Psi\rangle$ of H with respect to the tensor network state defined in (a); the expectation value equals to $\vec{\mathbf{A}}_i^\dagger \mathbf{H}_i \vec{\mathbf{A}}_i$ with $\mathbf{H}_i$ representing the effective Hamiltonian, drawn with dashed lines.

In the algorithmic approach to optimize the TTNS, one can use tools known in literature[50,51,219,228] and optimize the network site-by-site as in the DMRG. The fact that the tree tensor network does not contain any loops allows an exact mathematical treatment[81,136] (see in section 3.2). For $z = 2$, the DMRG algorithm is recovered. The TTNS algorithm is similar to a DMRG calculation with z blocks instead of two, where a block consists of all of the sites within one of the branches emerging from site i (see Fig. 28(a)).

As in DMRG, the TTNS algorithm consists in the variational optimization of the tensors $\mathbf{A}_i$ in such a way that the energy is minimized (with the constraint that the norm of the state remains constant). This is equivalent to minimizing the functional

$$F = \langle\Psi|\mathbf{H}|\Psi\rangle - E\left(\langle\Psi|\Psi\rangle - 1\right), \tag{179}$$

with $|\Psi\rangle = |\Psi(\mathbf{A}_1, \ldots, \mathbf{A}_d)\rangle$. This functional is non-convex with respect to all parameters $\{\mathbf{A}_1, \ldots, \mathbf{A}_d\}$. However, by fixing all tensors $\mathbf{A}_k$ except $\mathbf{A}_i$, due to the tensor network structure of the ansatz, it is quadratic in the parameters $\mathbf{A}_i$ associated with one lattice site i.

As depicted in Fig. 27(a), the tensor network state can be separated in two parts: the tensor $\mathbf{A}_i$ that is supposed to be optimized and an environment tensor $\mathcal{E}_i$ that is formed by contracting all tensors except $\mathbf{A}_i$. $\mathbf{A}_i$ is connected to the environment tensor $\mathcal{E}_i$ by z_i

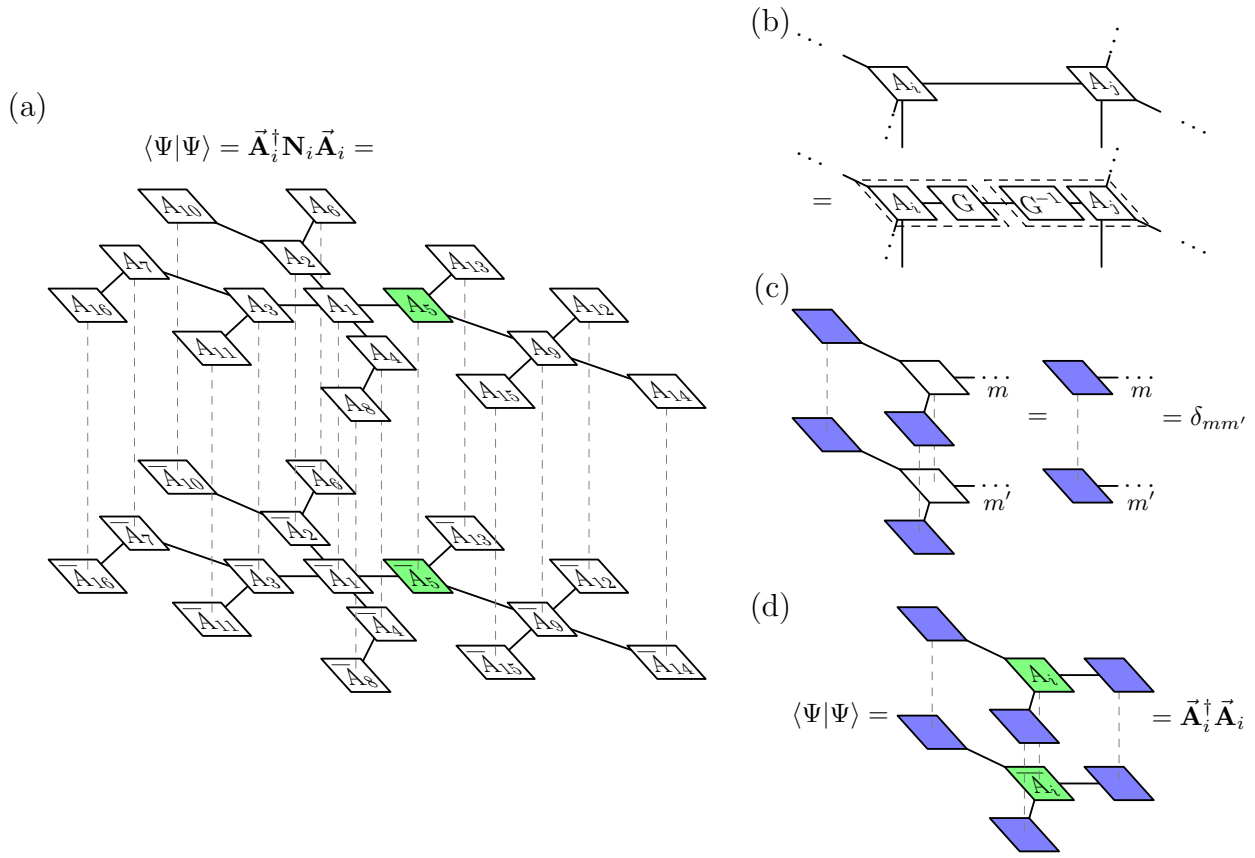


Figure 28: (a) norm $\langle\Psi|\Psi\rangle$ of the TTNS defined in Fig. 26; the tensor network picture of the norm corresponds to a two-layer structure, with the ket $|\Psi\rangle$ being on top and the bra $\langle\Psi|$ on bottom. For better readability, the contracted physical indices are drawn with dashed lines. (b) gauge transformation in a tensor network state: the state remains invariant if matrices G and G^{-1} are inserted at one bond and merged with the adjacent tensors. (c) orthonormalization condition imposed on all tensors (blue) except $\mathbf{A}_i$ (green). (d) norm of the TTNS with all tensors except $\mathbf{A}_i$ fulfilling the orthonormalization condition.

virtual bonds, with z_i being the coordination number of site i. Using this separation, it is evident that $\langle\Psi|\Psi\rangle = \vec{\mathbf{A}}_i^\dagger \mathbf{N}_i \vec{\mathbf{A}}_i$ and $\langle\Psi|\mathbf{H}|\Psi\rangle = \vec{\mathbf{A}}_i^\dagger \mathbf{H}_i \vec{\mathbf{A}}_i$, as shown in Fig. 27(b) and (c). $\vec{\mathbf{A}}_i$ is thereby the reshaped $q \times M \times \cdots \times M$-tensor $\mathbf{A}_i$ into a qM^{z_i}-dimensional vector. The inhomogenity $\mathbf{N}_i$ and the effective Hamiltonian $\mathbf{H}_i$ with respect to site i are matrices of size $qM^{z_i} \times qM^{z_i}$ that are obtained by contracting all tensors except $\mathbf{A}_i$ in the tensor expressions for $\langle\Psi|\Psi\rangle$ and $\langle\Psi|\mathbf{H}|\Psi\rangle$, respectively (see Fig. 27(b) and (c)).

The optimal parameters $\mathbf{A}_i$ can be found by minimizing the quadratic function

$$F(\vec{\mathbf{A}}_i) = \vec{\mathbf{A}}_i^\dagger \mathbf{H}_i \vec{\mathbf{A}}_i - E\left(\vec{\mathbf{A}}_i^\dagger \mathbf{N}_i \vec{\mathbf{A}}_i - 1\right), \tag{180}$$

which is equivalent to solving the generalized eigenvalue problem

$$\mathbf{H}_i \vec{\mathbf{A}}_i = E \mathbf{N}_i \vec{\mathbf{A}}_i. \tag{181}$$

For a network without loops, it is always possible to set $\mathbf{N}_i$ equal to the identity, which accounts for numerical stability because the generalized eigenvalue problem reduces to an

ordinary one. The reason for this possiblity is the gauge degree of freedom that exists in tensor networks: without changing the state, matrices G and G^{-1} can always be inserted at a bond and merged with the adjacent tensors, as depicted in Fig. 28(b). Because of this gauge degree of freedom, each tensor $\mathbf{A}_j$ for $j \neq i$ can be enforced to fulfill the orthonormalization condition

$$\sum_{\vec{m}_{\text{in}}} A_j(\alpha_j, \vec{m}_{\text{in}}, m_{\text{out}}) A_j(\alpha_j, \vec{m}_{\text{in}}, m'_{\text{out}}) = \delta_{m_{\text{out}} m'_{\text{out}}}. \tag{182}$$

Here, the "out"-index m_{out} is the index pointing towards site i, the remaining indices are denoted "in"-indices $\vec{m}_{\text{in}}$. In pictorial form, this condition is illustrated in Fig. 28(c). The mathematical operation that endows tensors with the orthonormalization condition is the QR-decomposition which is numerically stable[169]. Due to the orthonormalization condition, the tensor network for the norm of the TTNS, as shown in Fig. 28(a), can be "cropped" from the leaves towards site i, until only the tensors $\mathbf{A}_i$ and $\overline{\mathbf{A}_i}$ at site i remain. The norm of the TTNS then simplifies to $\langle\Psi|\Psi\rangle = \vec{\mathbf{A}}_i^\dagger \vec{\mathbf{A}}_i$, which makes $\mathbf{N}_i = 1$ (see Fig. 28(d)).

The challenge that remains is to calculate the effective Hamiltonian $\mathbf{H}_i$ of the eigenvalue problem. As mentioned before, it is obtained by contracting all tensors except $\mathbf{A}_i$ and $\overline{\mathbf{A}_i}$ in the tensor network of $\langle\Psi|\mathbf{H}|\Psi\rangle$. In case of a TTNS, the contraction is efficient if the Hamiltonian $\mathbf{H}$ is present in the form of a tree tensor network, as well. The tree network of the Hamiltonian shall have the same structure as the tree network of the state. In analogy to the definition of the TTNS, a tensor

$$h_i(\alpha'_i, \alpha_i, m_{i,1}, \ldots, m_{i,z_i}) \tag{183}$$

is associated to each site i with physical indices α'_i and α_i and virtual indices $m_{i,1}, \ldots, m_{i,z_i}$. The coefficients $\mathbf{H}(\alpha'_1, \ldots, \alpha'_d, \alpha_1, \ldots, \alpha_d)$ are then obtained by contracting the virtual indices of the tensors $\mathbf{h}_i$ according to the tree network. For $z = 2$, this corresponds to the representation of the Hamiltonian as a *Matrix Product Operator* (MPO), as depicted in Fig. 29(b1). For $z > 2$, this concept is generalized to a *Tree Tensor Network Operator* (TTNO), which is illustrated in Fig. 29(b2). In fact, for local Hamiltonians it is always possible to find a representation as an MPO or TTNO with *constant* dimension of the virtual bonds[191]. For non-local Hamiltonians of the form (43), as arising in quantum chemistry, it is always possible to find an MPO- or TTNO-form with bond-dimension $\mathcal{O}(d^2)$.

Once the Hamiltonian $\mathbf{H}$ is represented as TTNO with the same network structure as the TTNS $|\Psi\rangle$, the tensor network form of the expectation value corresponds to a three-layer object, as depicted in Fig. 29(a), with the ket $|\Psi\rangle$ consisting of component tensors $\mathbf{A}_i$ being on top, the bra $\langle\Psi|$ consisting of component tensors $\overline{\mathbf{A}_i}$ on bottom, and the Hamiltonian $\mathbf{H}$, represented as TTNO of component tensors $\mathbf{h}_i$ in the middle. By starting from the leaves and proceeding inwards towards site i, this network can be contracted efficiently (i.e. polynomially in d and M), yielding the expectation value $\langle\Psi|\mathbf{H}|\Psi\rangle$ and, if $\mathbf{A}_i$ and $\overline{\mathbf{A}_i}$ are omitted, the effective Hamiltonian $\mathbf{H}_i$. In order to reduce computational costs related to the diagonalization of the effective Hamiltonian a half-renormalization scheme has also been introduced[172].

For more detailed derivations we refer to the original papers[169,170,172].

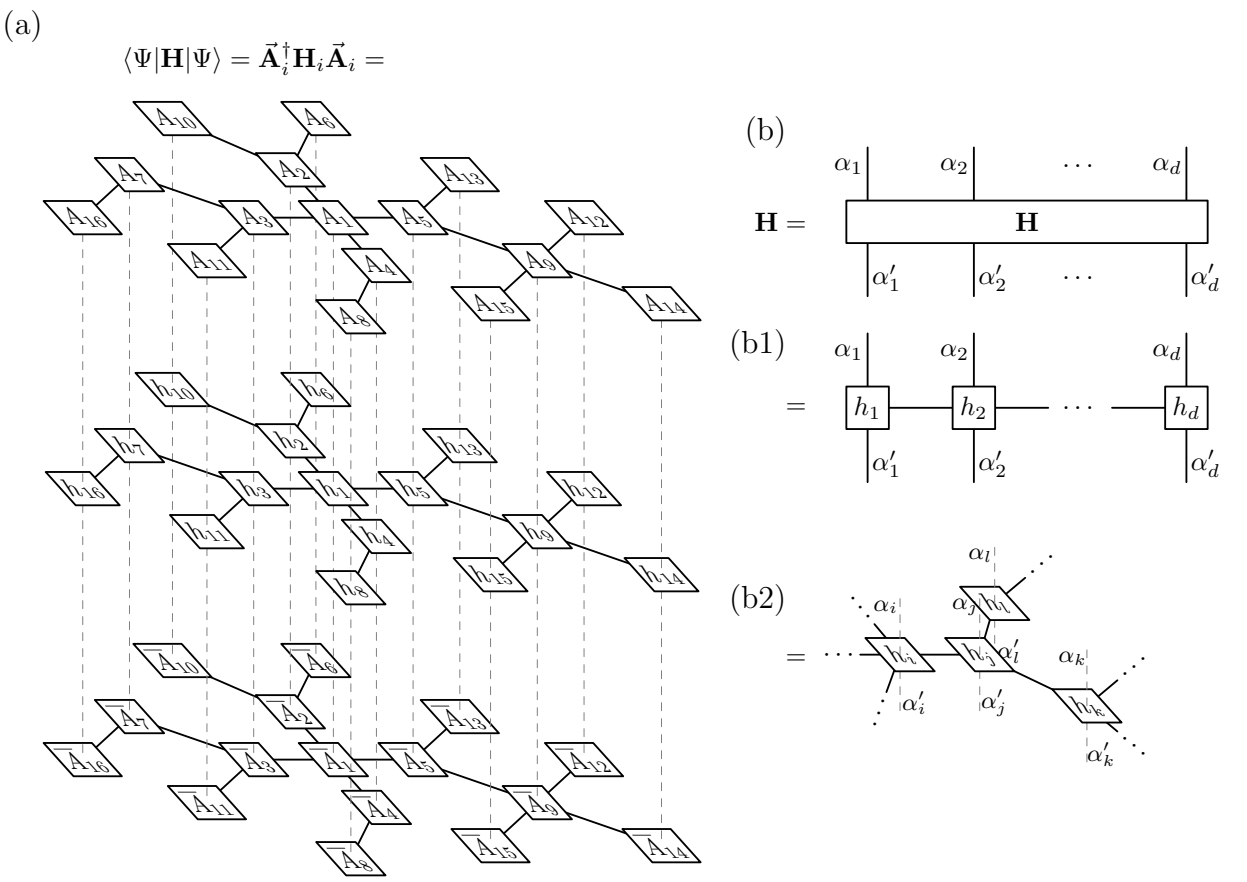


Figure 29: (a) expectation value $\langle\Psi|\mathbf{H}|\Psi\rangle$ with respect to the TTNS defined in Fig. 26; the tensor network picture of the expectation value corresponds to a three-layer structure, with the ket $|\Psi\rangle$ consisting of component tensors $\mathbf{A}_i$ being on top, the bra $\langle\Psi|$ consisting of component tensors $\overline{\mathbf{A}_i}$ on bottom, and the Hamiltonian $\mathbf{H}$, represented as TTNO of component tensors $\mathbf{h}_i$ in the middle. For better readability, the contracted physical indices are drawn with dashed lines. (b) decomposition of the Hamiltonian as MPO (b1) and TTNO (b2).

4.3.5 Efficient factorization of the interaction terms

When the d-orbital system is partitioned into several subsystems the Hamiltonian is built from terms acting *within* the subsystems and from terms *among* the subsystems. During the course of the iterative diagonalization of the effective Hamiltonian acting on the $M_L M_R$ dimensional subspace, the matrix vector multiplication $\mathbf{H}|\Psi\rangle$ is performed several times[55,118]. For a bipartite split using the matricization of U discussed in Sec. 3.4, $|\Psi\rangle$ is converted to a matrix with size $M_L \times M_R$ and the matrix vector multiplication is formed as two matrix-matrix multiplication of operator pairs as $\mathbf{X}_i^{(\mathrm{L})}\mathbf{U}(\mathbf{Y}_j^{(\mathrm{R})})^{\mathrm{T}}$ where $\mathbf{X}_i^{(\mathrm{L})}$ and $\mathbf{Y}_j^{(\mathrm{R})}$ are operators acting on the left and right subsystem, respectively.

In order to treat long-range interactions efficiently, the interaction terms must be factorized, thus the matrix and tensor algebra during the diagonalization procedure is simplified. This is called partial summation[266]. For example, considering a two-orbital interaction in

general, like the two-operator term in (43), for a bipartite split of the system

$$\mathbf{H}_{\text{two}} = \sum_{i,j\in(\text{L})} T_{ij}\mathbf{X}_i^{(\text{L})}\mathbf{Y}_j^{(\text{L})} + \sum_{\substack{i\in(\text{L})\\ j\in(\text{R})}} T_{ij}\mathbf{X}_i^{(\text{L})}\mathbf{Y}_j^{(\text{R})} + \sum_{\substack{i\in(\text{R})\\ j\in(\text{L})}} T_{ij}\mathbf{X}_i^{(\text{R})}\mathbf{Y}_j^{(\text{L})} + \sum_{i,j\in(\text{R})} T_{ij}\mathbf{X}_i^{(\text{R})}\mathbf{Y}_j^{(\text{R})}, \quad (184)$$

one of the coupling between the two subsystems (the second term above) can be simplifed as

$$\mathbf{H}^{(\text{L})(\text{R})} = \sum_{j\in(\text{R})} \left(\sum_{i\in(\text{L})} T_{ij}\mathbf{X}_i^{(\text{L})} \right) \mathbf{Y}_j^{(\text{R})} = \sum_{j\in(\text{R})} \mathcal{A}_j^{(\text{L})(\cdot)}\mathbf{Y}_j^{(\text{R})}. \quad (185)$$

Here $\mathcal{A}_j^{(\text{L})(\cdot)} = \sum_{i\in L} T_{ij}\mathbf{X}_i^{(\text{L})}$ is called *one-orbital auxiliary operator.*

Therefore, the number of operator multiplications reduces from d^2 to d. Symbolically this can be written in a compact form: we assign a label to each subsystem and form the total system by adding together the subsystems. This sum is raised to the power given by the number of operators corresponding to the given interaction. For example, for the four operator term, coming from the Coulomb intercation, $\mathbf{H}_{\text{four}} = \sum_{ijkl} V_{ijkl}\mathbf{X}_i\mathbf{Y}_j\mathbf{Z}_k\mathbf{W}_l$ in (43) and for the bipartite split (subsystems (L) and (R)), this can be factorized as $((\text{L})+(\text{R}))^4 = (\text{L})^4+4(\text{L})^3(\text{R})+6(\text{L})^2(\text{R})^2+4(\text{L})(\text{R})^3+(\text{R})^4$. Constant factors comes from the permutation of indices and exponents show the number of operators acting within the corresponding subsystem. Therefore, when the first three operators act on the (L) subsystem and the last operator on the (R) subsystem then

$$\mathbf{H}^{(\text{L})(\text{L})(\text{L})(\text{R})} = \sum_{l\in(\text{R})} \left(\sum_{i,j,k\in(\text{L})} V_{ijkl}\mathbf{X}_i^{(\text{L})}\mathbf{Y}_j^{(\text{L})}\mathbf{Z}_k^{(\text{L})} \right) \mathbf{W}_l^{(\text{R})} = \sum_{l\in(\text{R})} \mathcal{A}_l^{(\text{L})(\text{L})(\text{L})(\cdot)}\mathbf{W}_l^{(\text{R})}, \quad (186)$$

thus the number of operator multiplications reduces from d^4 to d by forming a three-orbital auxiliary operator $\mathcal{A}_l^{(\text{L})(\text{L})(\text{L})(\cdot)}$. Similarly, when the first two operators act on the (L) subsystem and the last two operators on the (R) subsystem then

$$\mathbf{H}^{(\text{L})(\text{L})(\text{R})(\text{R})} = \sum_{k,l\in(\text{R})} \left(\sum_{i,j\in(\text{L})} V_{ijkl}\mathbf{X}_i^{(\text{L})}\mathbf{Y}_j^{(\text{L})} \right) \mathbf{Z}_k^{(\text{R})}\mathbf{W}_l^{(\text{R})} = \sum_{k,l\in(\text{R})} \mathcal{A}_{kl}^{(\text{L})(\text{L})(\cdot)(\cdot)}\mathbf{Z}_k^{(\text{R})}\mathbf{W}_l^{(\text{R})} \quad (187)$$

thus the number of operator multiplications reduces from d^4 to d^2 by froming two-orbital auxiliary operators $\mathcal{A}_{kl}^{(\text{L})(\text{L})(\cdot)(\cdot)}$. Extensions for more subsystems used in QC-DMRG and QC-TTNS is straighforward. For example, for subsystems (l), $(l+1)$, $(l+2)$, (r), the two-operator term is composed from the following terms as $((\text{l})+(l+1)+(l+2)+(\text{r}))^2 = (\text{l})^2+2(\text{l})(l+1)+2(\text{l})(l+2)+2(\text{l})(\text{r})+(l+1)^2+2(l+1)(l+2)+2(l+1)(\text{r})+(l+2)^2+2(l+2)(\text{r})+(\text{r})^2$ and the four-operator term factorizes as $((\text{l})+(l+1)+(l+2)+(\text{r}))^4 = (\text{l})^4+4(\text{l})^3(l+1)+4(\text{l})^3(l+2)+4(\text{l})^3(\text{r})+6(\text{l})^2(l+1)^2+12(\text{l})^2(l+1)(l+2)+\ldots$ It is worth to note that symmetries of V_{ijkl} can be used to reduce the number of independent terms.

Renormalization of multi-orbital operators, i.e, when more than one operator act in the same subsystem, requires specal care since they cannot be calculated accurately as a product

of the renormalized operators. For example, if i, j belong to the same DMRG block due to the truncation of the Hilbert-space, $\mathbf{O}\mathbf{O}^\dagger \neq \mathbb{I}$.

$$\mathbf{O}(\mathbf{X}_i^{(\mathrm{L})}\mathbf{Y}_j^{(\mathrm{L})})\mathbf{O}^\dagger \neq \mathbf{O}\mathbf{X}_i^{(\mathrm{L})}\mathbf{O}^\dagger\mathbf{O}\mathbf{Y}_j^{(\mathrm{L})}\mathbf{O}^\dagger \tag{188}$$

Therefore, multi-orbital operators must be renormalized independently and stored. As an example, the renormalization of a four-orbital operator acting on the $(\mathrm{L}) = (\mathrm{l})\bullet$ composite system is $\mathbf{O}\mathcal{A}^{(\mathrm{L})(\mathrm{L})(\mathrm{L})(\mathrm{L})}\mathbf{O}^\dagger = \mathbf{O}(\sum_{ijkl\in(\mathrm{L})} V_{ijkl}\mathbf{X}_i^{(\mathrm{L})}\mathbf{Y}_j^{(\mathrm{L})}\mathbf{Z}_k^{(\mathrm{L})}\mathbf{W}_l^{(\mathrm{L})})\mathbf{O}^\dagger$, where the auxuiliary operator $\mathcal{A}^{(\mathrm{L})(\mathrm{L})(\mathrm{L})(\mathrm{L})}$ is decomposed into further auxiliary operators as follows

$$\begin{aligned}\mathcal{A}^{(\mathrm{L})(\mathrm{L})(\mathrm{L})(\mathrm{L})} &= \mathcal{A}^{(\mathrm{l})(\mathrm{l})(\mathrm{l})(\mathrm{l})} \otimes \mathbb{I}_{l+1} + \mathcal{A}_{l+1}^{(\mathrm{l})(\mathrm{l})(\mathrm{l})(\cdot)} \otimes \mathbf{W}_{l+1} + \mathcal{A}_{l+1,l+1}^{(\mathrm{l})(\mathrm{l})(\cdot)(\cdot)} \otimes \mathbf{Z}_{l+1}\mathbf{W}_{l+1} \\ &+ \mathcal{A}_{l+1}^{(\mathrm{l})(\cdot)(\cdot)(\cdot)} \otimes \mathbf{Y}_{l+1}\mathbf{Z}_{l+1}\mathbf{W}_{l+1} + \mathbb{I}^{(\mathrm{l})} \otimes V_{l+1,l+1,l+1,l+1}\mathbf{X}_{l+1}\mathbf{Y}_{l+1}\mathbf{Z}_{l+1}\mathbf{W}_{l+1}\end{aligned} \tag{189}$$

In summary, the numerical effort of the QC-DMRG and QC-TTNS algorithms has two major contributions. On the one hand, the number of block states is crucial: The numerical effort for calculating one term of the effective Hamiltonian by tensor contraction scales as M^{z+1} for trees of arbitrary coordination number z. On the other hand, this calculation has to be performed for each term in the Hamiltonian, and using the summation tricks as described above the scaling is d^2M^{z+1}. Since $\mathcal{O}(d)$ iteration steps are required for convergence, the overall time of the algorithms scale as d^3M^{z+1}.

4.4 Optimization of convergence properties

In order to use QC-DMRG and QC-TTNS as black box methods, it is mandatory to utilize various concepts inherited from quantum information theory[15,70,127,129,169,170,201]. In this section we briefly discuss some entanglement based optimization procedures which are used to minimize the *overall entanglement*, expressed as a cost function[15,201],

$$I_{\mathrm{overall}} = \sum_{i,j} I_{ij} d_{ij}^{\eta}. \tag{190}$$

Here d_{ij} is the *distance function* between orbital i and j, in the graph-theoretical sense, I_{ij} is the *two-orbital mutual information* given in (164) and η is some exponent. Therefore, the correlations between the pairs of orbitals is weighted by the distance d_{ij}. The distance d_{ij} depends on the tensor topology, and it is defined as the length of the shortest path connencting i and j in the tensor network. In the special case of MPS, the distance is simply $d_{ij} = |i - j|$.

The physical motivation behind the quantity I_{overall} is that in a given iteration step the Schmidt rank is related to the number and strength of the entanglement bonds between the left and right blocks, thus if two highly correlated orbitals are located far from each other then they give a large contribution until they fall into the same block. Since the overall cost is related to the sum of the Schmidt ranks, the major aim is to reduce the ranks for each iteration steps. The optimization methods surveyed in this section serve for the manipulation of this cost function I_{overall} in three different ways: by changing d_{ij} by reordering the component tensors for a given tensor topology (section 4.4.4); by changing d_{ij} by altering the tensor topology itself (section 4.4.5); or by changing I_{ij} by transforming

the orbital basis (section 4.4.6). Besides this, there are other factors which effects the convergence rate and computational time: using Dynamical Block State Selection (DBSS) methods (section 4.4.3); using entanglement based network initialization (section 4.4.7); or reducing the Hilbert space by taking symmetries into consideration (section 4.4.8).

4.4.1 Error sources and data sparse representation of the wavefunction

As has been discussed before, the success and numerical efficiency of the QC-DMRG and QC-TTNS algorithm rely on a subsequent application of the singular value decomposition[136,169,172,212] (section 4.2.1) while the performance depends on the level of entanglement encoded in the wave function[127,241]. In each DMRG (or TTNS) step, the basis states of the system block are then transformed to a new *truncated basis* set by a unitary transformation based on the preceeding SVD[211]. This transformation depends therefore on how accurately the environment is represented[132,163] as well as on the level of truncation[125]. As a consequence the accuracy of the DMRG method is governed by the truncation error, $\delta\varepsilon_{\rm TR}$, as well as by the environmental error, $\delta\varepsilon_{\rm sweep}$[132]. The latter is minimized in each DMRG sweep (macro-iteration) by a successive application of the SVD going through the system back and forth. Since $\dim(\Xi_{\rm DMRG}) \ll \dim(\Lambda_{\rm FCI})$ DMRG provides a data-sparse representation of the wavefunction, thus the sparsity can be defined as $\dim(\Xi_{\rm DMRG})/\dim(\Lambda_{\rm FCI})$ for a given error margin.

As an example, relevant quantities as a function of DMRG iteration steps are shown in Fig. 30 for LiF at $d_{\text{Li-F}} = 3.05$ a.u. for two different tensor arrangements (ordering). Since DMRG is a variational method it converges to the full-CI energy from above as is apparent in the top panels of Fig 30. Close to the turning points when either left or right block contains a single orbital M_l or M_r drops to $4 = q$. Although the truncation error fluctuates between 10^{-16} and 10^{-6} for both tensor arrangements (ordering) and the size of the superblock Hilbert space is at most 400, a much lower energy has been reached with the optimized ordering. This clearly shows that in order to minimize $\delta\varepsilon_{\rm sweep}$ and avoid DMRG to converge to a local minima besides sweeping the tensor arrangement must also be optimized as will be discussed below.

Using an optimized ordering the convergence of the ground state energy for LiF at $d_{\text{Li-F}} = 3.05$ a.u. as a function of DMRG sweepings for various fixed number of block states is shown in Fig 31(a). Taking the limit of zero energy change between two sweeps $E(M, \delta\varepsilon_{\rm sweep} = 0)$ for a given M and assuming $M_l = M_r = M$ various extraplotaion schemes as a function of M have been introduced[16,39,125,150,158] in order to provide a good estimate for the truncation-free solution. A more rigorous extrapolation scheme is based on the truncation error[132], i.e., once the environmental error is eliminated, the relative error, $\Delta E_{\rm rel} = (E_{\rm DMRG} - E_{\rm FCI})/E_{\rm FCI}$, is determined by $\delta\varepsilon_{\rm TR}$ as

$$\ln \Delta E_{\rm rel} = a \ln \delta\varepsilon_{\rm TR} + b\,. \tag{191}$$

When the number of block states are kept fixed the truncation error fluctuates within a full sweep (see Fig. 30) thus the largest truncation error within a full sweep determines the overall accuracy. In Fig. 31(d) the relative error of the ground state energy is shown as a function of the largest truncation error within the last full sweep on a log-log scale. The linear behavior allows one to obtain the truncation free energy by taking all the datapoints obtained upto a given $\delta\varepsilon_{\rm TR}$ and letting $E_{\rm FCI}$ as a free parameter denoted as E.

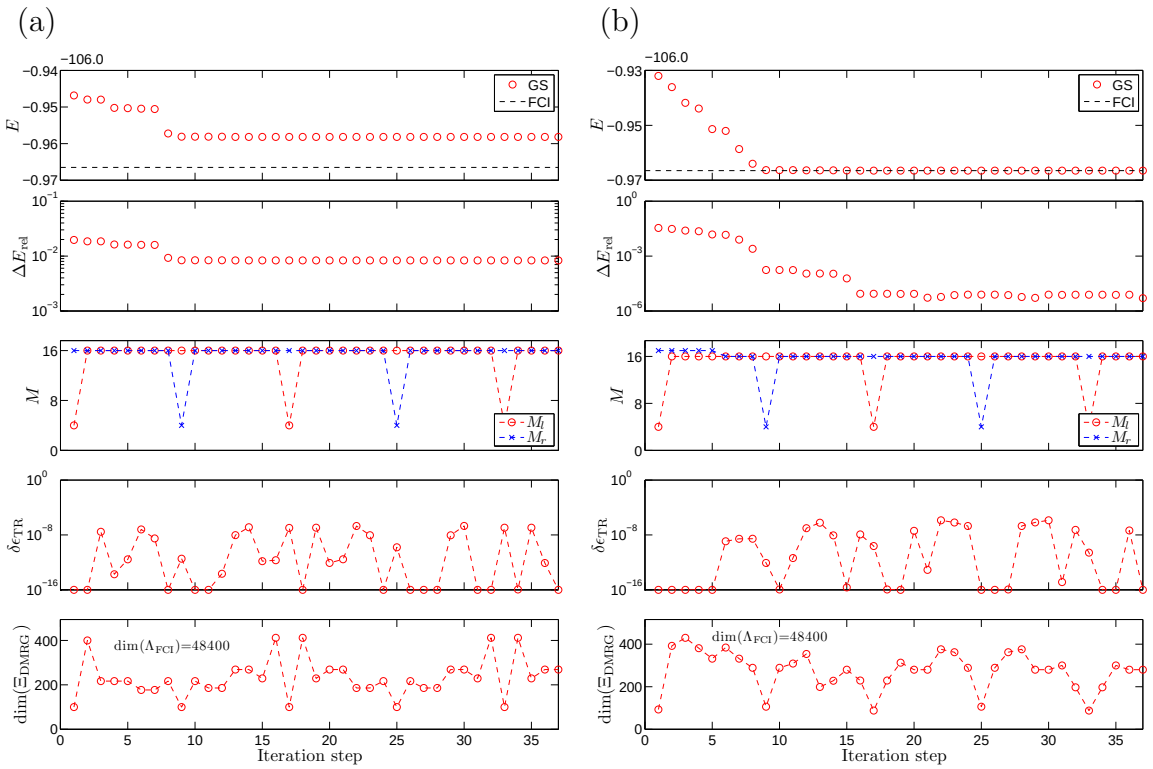


Figure 30: Ground state energy (E) in a.u., relative error ($\Delta E_{\rm rel} = (E_{\rm DMRG} - E_{\rm FCI})/E_{\rm FCI}$), number of block states (M_l, M_r), truncation error $\delta\varepsilon_{\rm TR}$, dimension of the superblock Hilbert space ($\Xi_{\rm DMRG}$), are shown as a function of DMRG itertaion steps for LiF at $d_{\text{Li-F}} = 3.05$ a.u. with CAS(6,12) with fixed $M_l = M_r = 16$ for a non-optimized tensor hierarchy (ordering) (a) and for an optimized tensor tensor hierarchy (ordering) (b).

4.4.2 Targeting several states together

As it is possible to calculate several lowest lying eigenstates of the superblock Hamiltonian using the Davidson[55] or Lánczos[118] algorithm, more eigenstates can be targeted within a single QC-DMRG or QC-TTNS calculation[57,74,125,126,139,165,170,173,218,264].

In this case the total system is no longer treated as a pure state but as a mixed state with mixing weights $p_\gamma > 0$ (with $\gamma = 1, \ldots, n$ and $\sum_\gamma p_\gamma = 1$), the reduced subsystem density matrix can be formed from the reduced density matrices ρ_γ of the lowest n eigenstates $|\Psi_\gamma\rangle$ as $\rho = \sum_\gamma p_\gamma \rho_\gamma$. The optimal choice of the p_γ distribution, however, is not established yet. As an example, energies of the ground state and first excited state obtained for the LiF at $d_{\text{Li-F}} = 3.05$ a.u. is shown in Fig. 32(a). It is worth mentioning that target states can also be formed based on the action of a given operator, i.e, besides the ground and excited states one can inlcude states by applying a given operator to the ground state. For more details we refer to the literature[178].

For multi-target states with equal weights $p_\gamma \equiv p = 1/n$, we minimize the sum $\sum_{\gamma=1}^{n} \langle\Psi_\gamma|H|\Psi_\gamma\rangle$ constrained to the orthogonality condition $\langle\Psi_\beta|\Psi_\gamma\rangle = \delta_{\beta,\gamma}$. Clearly the minimum of this functional is the sum of the n lowest eigenvalues $E_0 + \cdots + E_{n-1}$ of the Hamiltonian H, and a minimizer is provided by the first n eigenfunctions. In an MPS framework, the tensor $U(\alpha_1, \ldots, \alpha_d, \gamma)$ corresponding to the γ^{th} eigenstate with order $d+1$ as is shown in Fig. 33(a)

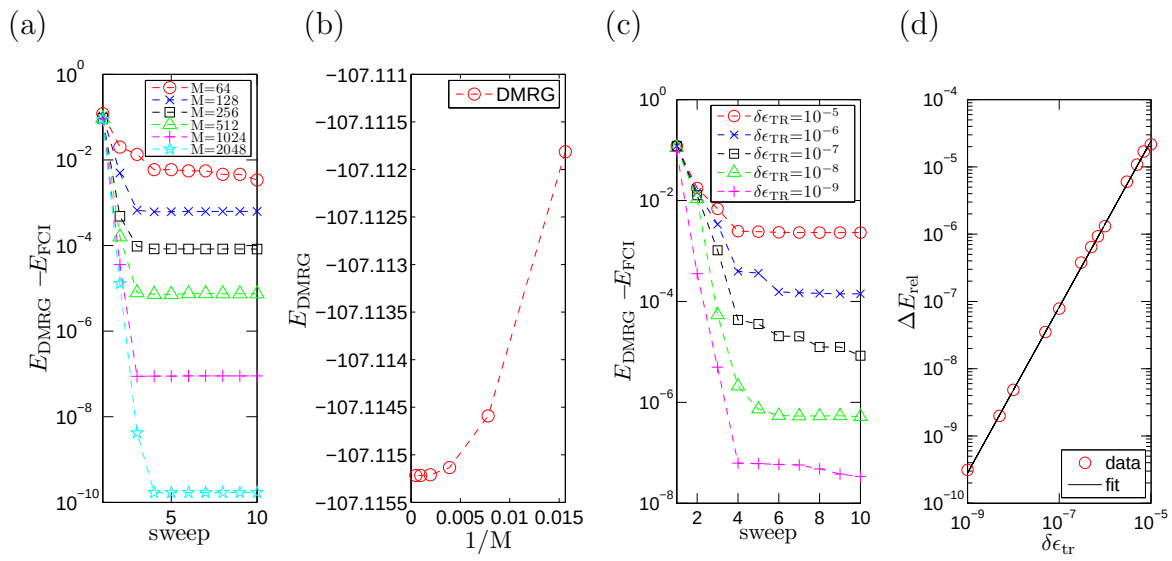


Figure 31: Convergence of the ground state energy for LiF at $d_{\text{Li-F}} = 3.05$ a.u. as a function of DMRG sweeping for various fixed number of block states (a), $E(M, \delta\varepsilon_{\text{sweep}} = 0)$ as a function of $1/M$(b), ground state energy as a function of DMRG sweeping for various fixed $\delta\varepsilon_{\text{TR}}$ using DBSS procedure (c), and $\Delta E_{\text{rel}}(\delta\varepsilon_{\text{TR}}, \delta\varepsilon_{\text{sweep}} = 0)$ as a function of $\delta\varepsilon_{\text{TR}}$ on a log-log scale (d). The solid lines are our fits.

can be expressed as a network shown in Fig. 33(b). Therefore, the network contains $d + 1$ component tensors and in each optimization step γ is shifted through the newtwork. Although, this procedure is commonly used in the DMRG community, it is worth mentioning that γ index has a different physical meaning than the α indices. Quite recently alternative methods to calculate excited states have also been itroduced [88,89,190,263].

4.4.3 Optimization of the Schmidt ranks using dynamic block state selection (DBSS) approach and entropy sum rule

The two-orbital variant of the DMRG method has originally been employed with a fixed number of block states as shown above while the degree of entanglement between the DMRG blocks for a given superblock configuration is related to the Schmidt rank r_{Sch} as discussed in Secs. 3 and 4.3.3. Therefore, the fluctuation of the truncation error makes the utilization of Eq. (191) less stable. It is more efficient to control the truncation error $\delta\varepsilon_{\text{TR}}$ at each renormalization step and change the number of block states dynamically [125].

Alternatively, one can control the truncation in terms of the quantum information loss χ, expressed by the von Neumann and Rényi entropies [129]. In a given DMRG renormalization step denoting by $S^{(\text{l})}$ the entropy of the left block of length l and by S_{l+1} the entropy of the $l + 1^{\text{th}}$ orbital, the sum of the entropies of these subsystems is reduced by forming a larger block, $(\text{L}) \equiv (\text{l})\bullet$, is given as

$$S^{(\text{l})} + S_{l+1} - S^{(\text{L})} = I^{(\text{l})} \geq 0, \tag{192}$$

where the mutual information $I^{(\text{l})}$ quantifies the correlation between the subsystem and the orbital (similarly to the mutual information in Eq. (164), doing the same for two orbitals).

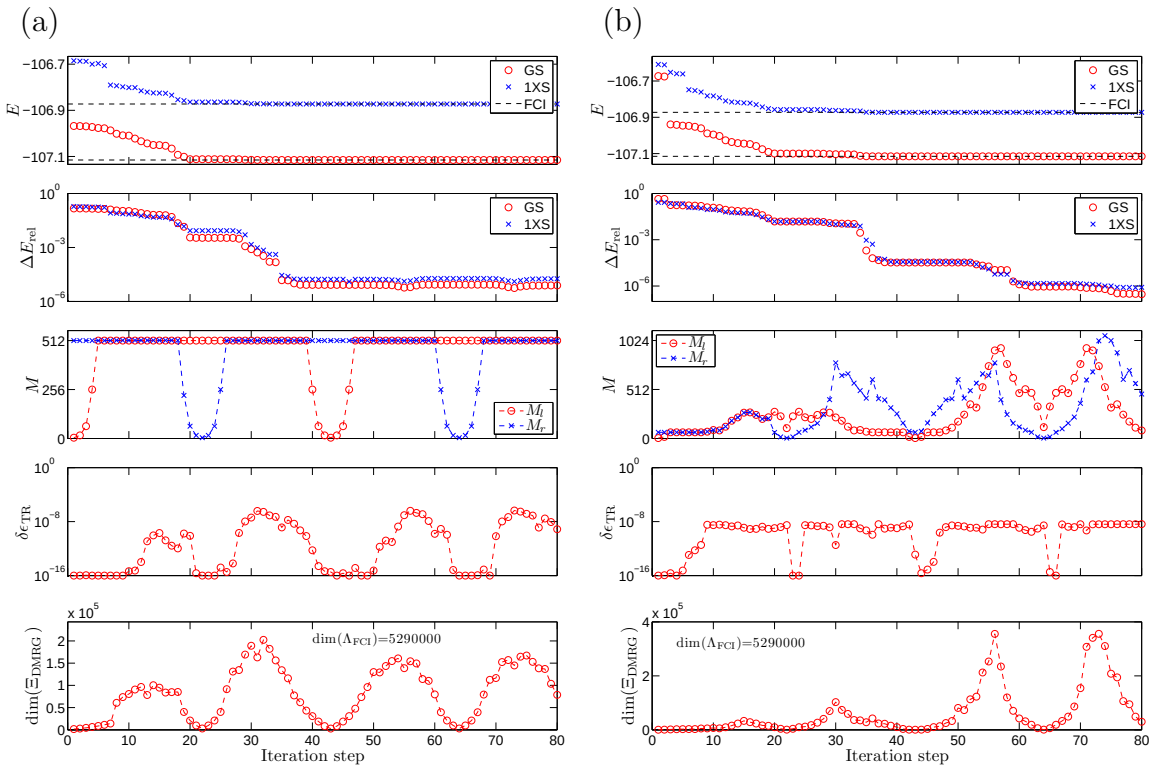


Figure 32: (a) Similar to Fig. 30(b) but for LiF at $d_{\text{Li-F}} = 3.05$ a.u. with CAS(6,25) and targeting the ground state and excited states within a single DMRG calculation with $M_l = M_r = 64$ block states, for optimized tensor hierarchy (ordering). (b) Similar to (a) but using the DBSS procedure with $M_{\min} = 64$ and $\chi = 10^{-7}$.

This means that if $I^{(\text{l})} > 0$ then we need more information for the description of the state of the (l) block and the $\bullet$ separatedly than for the description of them as a whole (L) = (l)$\bullet$, that is, they are correlated. A similar relation holds for the right block, (R) $\equiv \bullet$(r), as well. If an effective system of length $d+2$ is formed by adding two non-interacting orbitals to the right and left ends of the chain, all blocks containing 1 to d orbitals of the original system can be formed by the forward and backward sweeps. The total information gain during a half sweep can be calculated as $\sum_{l=1}^{d-1} I^{(\text{l})}$. In general, $I^{(\text{l})}$ is also a function of subsequent sweeps. However, once the DMRG method has converged, subsequent DMRG sweeps do not change $S^{(\text{l})}$ and S_l. If, additionally, all $M_l = q^l$ and $M_r = q^r$ basis states of the blocks are kept at each iteration step, i.e., no truncation is applied, a sum rule holds, which relates the total information gain within a full half sweep and the sum of orbital entropies given as

$$\sum_{l=1}^{d-1} I^{(\text{l})} = \sum_{l=1}^{d} S_l \,, \tag{193}$$

where we have used $S^{(1)} = S_1$ and $S^{(d)} = 0$.

This equality, however, does not hold in practical DMRG calculations since during the renormalization process $S^{(\text{L})}$ is reduced to $S^{(\text{L})}_{\text{Trunc}}$ due to the truncation of the basis states. Once the DMRG method has converged, the following equality should hold to a good accu-

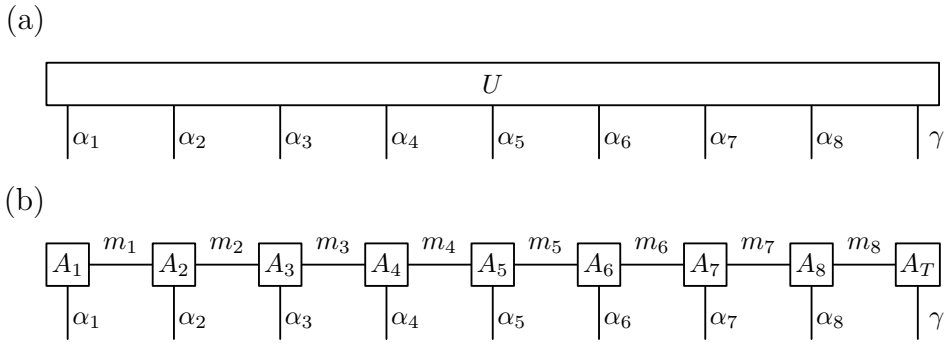


Figure 33: (a) Graphical representation of the full tensor $U(\alpha_1, \ldots, \alpha_d, \gamma)$ when excited states are also calculated. α stands for the physical indicies while γ labels the excited states. (b) Tensor network representation of the full tensor.

racy

$$\sum_{l=1}^{d-1} I^{(l)} \simeq \sum_{l=1}^{d} S_l - \sum_{l=1}^{d-1} \left(S^{(\mathrm{L})} - S^{(\mathrm{L})}_{\mathrm{Trunc}} \right) . \tag{194}$$

An analogous relationship holds for the backward sweep as well. In order to control the quantum information loss, M_L (or M_R) is increased systematically at each renormalization step until the following condition holds

$$S^{(\mathrm{L})} - S^{(\mathrm{L})}_{\mathrm{Trunc}} < \chi \,, \tag{195}$$

where χ is an a priori defined error margin. For $S^{(\mathrm{L})}$, i.e., before the truncation, $M_L = M_l q$ while for $S^{(\mathrm{L})}_{\mathrm{Trunc}}$ according to Eq. (195) $M_L^{\mathrm{Trunc}} \leq M_l q$ is used. This approach guarantees that the number of block states are adjusted according to the entanglement between the DMRG blocks and the a priori defined accuracy can be reached. In addition, an entropy sum rule based on Eq. (194) can be used as an alternative test of convergence[129].

In order to reduce the possibility of convergence to a local minima the minimum number of block states, M_{min} must also be introduced. Setting $M_{\mathrm{min}} \simeq q^3$ or q^4 is sufficient in most cases. The maximum number of block states selected dynamically during the course of iterations denoted by M_{max} determines wheter a calculation for a given accuracy can be performed on the available computational resources. It is worth to emphasize that this approach does not work for the one-orbital variant of the DMRG algorithm since the Schmidt number of a one-orbital superblock configuration $M_L = M_l q$ cannot be larger than M_r. This prevents M_l to increase above M_r according to Eq. (154).

As an example, relevant quantities as a function of DMRG iteration steps are shown in Fig. 32(b) for LiF at $d_{\text{Li-F}} = 3.05$ a.u. for the optimized tensor arrangements (ordering) using the DBSS procedure with $M_{\mathrm{min}} = 64$ and $\chi = 10^{-7}$. In Fig. 31(c) the convergence of the ground state energy as a function of DMRG sweeping for various fixed $\delta\varepsilon_{\mathrm{TR}}$ using the DBSS procedure is shown. Using Eq. (191) and data points obtained for $\delta\varepsilon_{\mathrm{TR}} \geq 10^{-6}$ after the 10^{th} sweep the extrapolated energy is $E = -107.11519(2)$, for $\delta\varepsilon_{\mathrm{TR}} \geq 10^{-9}$ it is $E = -107.115216925(2)$, while $E_{\mathrm{exact}} = -107.1152169273$.

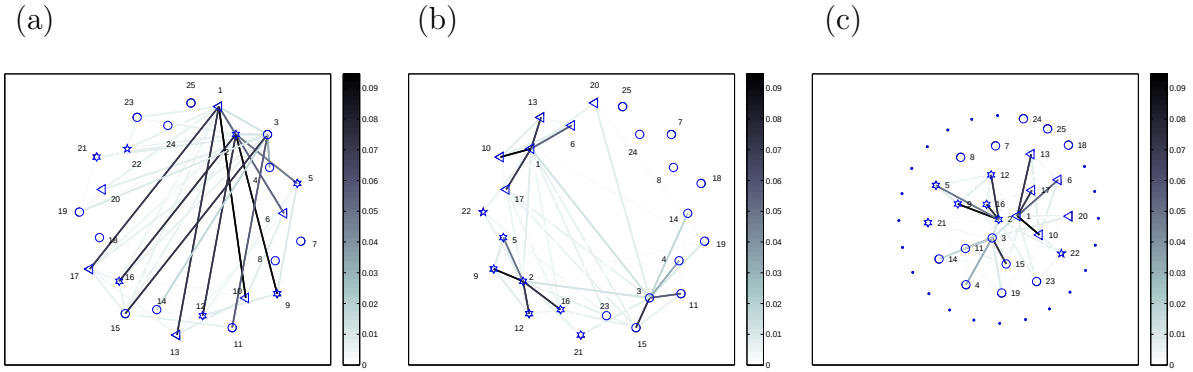


Figure 34: Optimization of tensor hierarchy (ordering) and topology by minimizing the overall entanglement I_{overall} for the LiF at the equilibrium bond length $r = 3.05$ a.u. (a) and (b) are for the one dimensional MPS like topology for the original ordering and for the optimized ordering, respectively. (c) Shows the optimized topology on the tree (small dots indicate not used grid points of the tree). The total quantum information I_{tot} given in (163) does not change but the overall entanglement I_{overall} drops significantly.

4.4.4 Optimization of the network hierarchy (ordering) and entanglement localization

As was briefly mentioned before, in order to use QC-DMRG as a black box method, first the arrangement of orbitals along a one-dimensional topology has to be optimized (ordering) in order to reduce the set of Schmidt ranks when the system is systematically partitioned into a left and right parts during the DMRG sweeping procedure [15,30,39,114,125–127,146,158,161,162,201,264,269]. This allows us to carry out calculations with much smaller number of block states using the DBSS approach [15,125,129] (section 4.4.3). For the one-dimensional tensor topology, i.e., for DMRG and MPS, the distance function is $d_{ij} = |i - j|$ in Eq. (190) and using $\eta = 2$ has the advantage that this optimization task can be carried out using concepts of spectral graph theory [14]. It follows that the so called *Fiedler vector* $x = (x_1, \dots x_d)$ is the solution that minimizes $F(x) = x^\dagger L x = \sum_{i,j} I_{ij}(x_i - x_j)^2$ subject to the constraints $\sum_i x_i = 0$ and $\sum_i x_i^2 = 1$, where the graph Laplacian is $L_{ij} = D_{ij} - I_{ij}$ with the diagonal $D_{ij} = \delta_{ij} \sum_{j'} I_{ij'}$. The second eigenvector of the Laplacian is the Fiedler vector [71,72] which defines a (1-dimensional) embedding of the graph on a line that tries to respect the highest entries of I_{ij} and the edge length of the graph. Ordering the entries of the Fiedler vector by non-increasing or non-decreasing way provides us a possible ordering. Usually the best ordering obtained with small number of block states also provide almost the best ordering for calculation performed with large number of block states, thus this task can be carried out with a limited number of block states. As an example, non-optimal and optimized tensor orderings for LiF at the equilibrium bond length $r = 3.05$ a.u. are shown in Figs. 34(a) and (b) for the one-dimensional network topology, respectively. For both tensor topologies $I_{\text{tot}} = 1.32$, given in Eq. (163), does not change but the overall entanglement I_{overall}, given in Eq. (190), drops significantly from 126.47 to 19.63. As a consequence, the maximum height and the spread of the block entropy is reduced significantly as shown in Fig. 15(a) and (b). Since Schmidt ranks are related to the block entropy, the same accurracy can be reached using much less block states and sweeps in the optimized (ordered) case. This leads to a huge save in CPU time and

memory[15,70,125,127].

4.4.5 Optimization of the network topology

Another possibility to minimize the overall entanglement I_{overall} given by Eq. (190) is to carry out network topology optimization. Based on the two-dimensional entanglement graph shown in Fig. 18(a), it is clear that orbitals are entangled with each other with different strengths. Therefore, when a tensor network is formed, the obvious choice is to allow the coordination number z_i to vary from orbital to orbital[169].

For the tree topology, see in section 4.3.4, d_{ij} in Eq. (190) can be computed as the distance from the center to i, plus the distance from the center to j, minus twice the distance from the center to their lowest common ancestor. The lowest common ancestor can be obtained within a linear preprocessing time $\mathcal{O}(d)$ and a constant query time using the Berkman's algorithm[24].

In practice, the optimal structure of the tree tensor network can be determined in a self-consistent way. First the one-orbital entropy and two-orbital mutual information is calculated with $z_i = 2$ and fixed small number of block states using the ordering of orbitals for which the T_{ij} and V_{ijkl} integral files were generated in order to determine entropy profiles qualitatively. Next orbitals with largest entropy values are placed close to the center of the network by keeping together those orbitals which are connected by large I_{ij} bonds as is shown in Fig. 34(c). Using such an optimized tensor topology the overall entanglement optimized for the $z_i = 2$ case can drop even further. In the present example for the LiF it reduces from $I^{\text{MPS}}_{\text{overall}} = 19.63$ to $I^{\text{TTNS}}_{\text{overall}} = 5.53$. As a result, the same numerical accuracy obtained with an MPS topology could have been reached with smaller number of block states and using less iteration steps when the optimized tree topology was used[170].

The overall efficiency of the QC-TTNS method is determined by two major parameters. On the one hand tensor ranks M decrease by going from QC-DMRG to QC-TTNS, but on the order hand the orders z of the tensors increases. Although, the computational cost of one iteration step is proportional to M^{z+1}, the number of tensors with $z = 1$ lying on the boundaries of the network increases exponentially when larger and larger systems are considered. Therefore, there is an expected crossover in cpu time between the full sweep of the QC-MPS and QC-TTNS. It is worth mentioning, that a two-orbital variant of the TTNS ansatz has also been considered in which the $z - 1$ environment blocks are mapped into one environment block through the so-called *half-renormalization* (HR) algorithm[172]. At present, optimization tasks are less established and straightforward, thus further developments are mandatory in order to fully utilize the potentials relying behind the TTNS algorithm.

4.4.6 Optimization of the basis using entanglement protocols

In the past 15 years various orbital bases have been employed to study quantum chemical systems[15,26,30,74,112,126,144,146,157,201,253,264,268]. Although the impact of a given basis on the efficiency of the QC-DMRG or QC-TTNS can be monitored by the convergence of the energy, a rigorous analysis in terms of the resulting entanglement patterns is mandatory in order to choose the most appropriate basis[70]. This is due to the fact, that the mutual information is orbital basis dependent. Therefore, besides orbital ordering and optimization of tensor

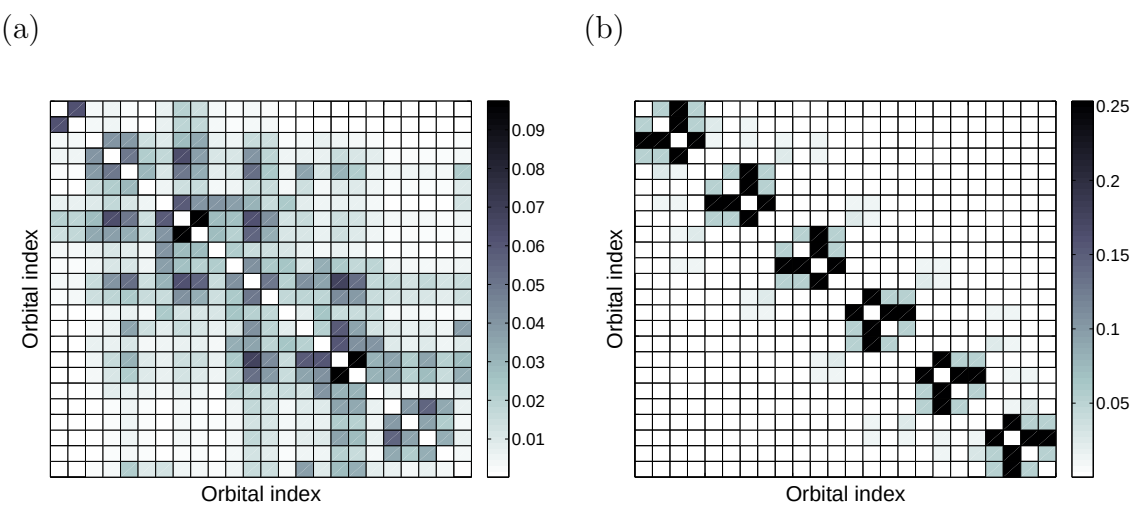


Figure 35: Colorscaled plot of two-orbital mutual information (for optimized orbital ordering using the Fiedler vector) for the ground state for Be_6 for a stretched structure, $d_{Be-Be} = 3.30$Å, using the DMRG method with canonical (a) and local (b) orbitals. $I_{tot} = 7.81$, $I_{overall} = 332.38$ with the canonical basis and $I_{tot} = 5.83$, $I_{overall} = 58.1$ with the local basis.

topology the overall entanglement $I_{overall}$ can be manipulated by changing the orbital basis as well. The performance of QC-DMRG and QC-TTNS can be optimized by using proper choice of the orbital basis, *i*.e., the same state can be obtained with much smaller number of block states[70,133,261]. As an example, entanglement patterns reported[70] for a ring of Be atoms using canonical HF and localized (Foster-Boys[31]) orbitals are shown in Fig. 35. The overall entanglement has been found to be much smaller in the latter case and as a consequence the same accuracy has been reached with much smaller number of block states.

Therefore, a main goal is to find a basis in which entanglement is localized as much as possible at the orbitals of the network, what would guarantee that a given precision could be attained with a smaller number of block states, and thus with less computational effort. One possibility is to find the optimal basis can be obtained by a canonical transformation of the fermionic modes using an $d \times d$ unitary matrix U, see section 2.5. In general, there are two ways to implement the basis transformation: one is based on the state and the other is based on the Hamiltonian, i.e.,

$$E(U) \equiv \langle\Psi|UHU^\dagger|\Psi\rangle = \langle\Psi(U)|H|\Psi(U)\rangle = \langle\Psi|H(U)|\Psi\rangle.$$

Since $E(U)$ is a non-convex function of the parameters U, it is a highly non-trivial problem to find the absolute minimum. Gradient search has been applied[169] to the function $E(U)$ expressed as

$$E(U) = \sum_{ij} \tilde{T}(U)_{ij}\langle a_i^\dagger a_j\rangle + \sum_{ijkl} \tilde{V}(U)_{ijkl}\langle a_i^\dagger a_j^\dagger a_k a_l\rangle$$

with $\tilde{T}(U) = UTU^\dagger$ and $\tilde{V}(U) = (U \otimes U)V(U \otimes U)^\dagger$, see equation (48). In this case, the correlation functions $\langle a_i^\dagger a_j\rangle$ and $\langle a_i^\dagger a_j^\dagger a_k a_l\rangle$ could be calculated with respect to the original state since they are independet of the parameters in U. The function $E(U)$ in this form and its gradient could be calculated explicitly and efficiently for different parameter sets U,

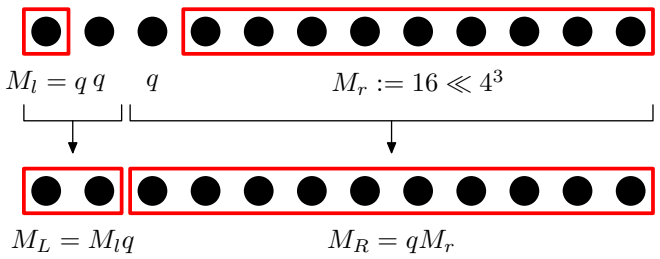


Figure 36: A superblock configuration with $l = 1$ and $r = 9$.

which made the gradient search feasible. This assures that the energy decreases significantly in the course of the algorithm since the orbital optimization is performed repeatedly during the course of the network optimization.

4.4.7 Optimization of the network initialization based on entanglement

Besides ordering, network topology and basis states optimization the optimal performance of SVD based methods is strongly effected by the initial conditions, or in other words by the initial matrix and tensor configurations. If a poorly approximated starting configuration is used, the convergence can be very slow and the DMRG can even be trapped in local minima[125,127,163]. In the past decade various solutions have been introduced in order to optimize network initialization[15,39,127,159,163,264]. In the following we focus on an entanglement based procedure.

Having a tensor network with a given topology and hierarchy (ordering) the elements of the component tensors are random numbers in the first ietartion step. In QC-DMRG and QC-TTNS methods various Ξ truncated Hilbert spaces can be formed from different subsets of the corresponding basis states in order to approximate Λ Hilbert space. In other words, for a given partitioning of the system into blocks various environment blocks can be generated for a given system block.

In case of the two-orbital QC-DMRG the optimization starts with a superblock configuration $l = 1$ and $r = d - 3$ as shown in Fig. 25. When the SVD is performed, the eigenvalue spectrum of the reduced density matrix of the (L) block depends on how the truncated basis was formed for the right block. Since the exact representation of the right block would require $M_r = q^{d-l-2}$ states, which is too large for large d, only a subset of orbitals is included to form the active space. As an example, three different environment blocks formed from three different subsets of $M_r = 16$ basis states (or $M_r = 17$ due to spin reflection symmetry) obtained for the superblock configuration $l = 1$ and $r = 9$ (see Fig. 36) of the LiF molecule for $d_{\text{Li-F}} = 3.05$ a.u. with CAS(6,12) are shown in table 6. Using an ordering according to the energy, the first, second and third orbitals are the Hartree-Fock (HF) orbitals. The selected $M_r = 16$ environment states together with the $M_l q^2 = 64$ states of the $(l) \bullet \bullet$ composite system fulfill the conservation of total number of particles with up and down spins, i.e., $N_\downarrow^{(\mathrm{l})} + N_{l+1,\downarrow} + N_{l+2,\downarrow} + N_\downarrow^{(\mathrm{r})} = 3$ and $N_\uparrow^{(\mathrm{l})} + N_{l+1,\uparrow} + N_{l+2,\uparrow} + N_\uparrow^{(\mathrm{r})} = 3$. By forming the bi-partite spliting of the system with $L = 2$ and $R = 10$ the eigenvalue spectrum of $\rho^{(\mathrm{L})}$ (and $\rho^{(\mathrm{R})}$) corresponding to the three subsets and the one corresponding to the exact solution obtained by $M_r = 8000$ block states are shown in Fig. 37. It is obvious that the block entropy

α_r	configuration 1									CI
1	–	–	–	–	–	–	–	–	–	0
2	–	–	–	–	–	–	–	–	↓	1
3	–	–	–	–	–	–	–	–	↑	1
4	–	–	–	–	–	–	–	–	↑↓	2
5	–	–	–	–	–	–	–	↓	–	1
6	–	–	–	–	–	–	–	↑	–	1
7	–	–	–	–	–	–	–	↓	↓	2
8	–	–	–	–	–	–	–	↑	↑	2
9	–	–	–	–	–	–	–	↓	↑	2
10	–	–	–	–	–	–	–	↑	↓	2
11	–	–	–	–	–	–	–	↑↓	–	2
12	–	–	–	–	–	–	↓	–	–	1
13	–	–	–	–	–	–	↑	–	–	1
14	–	–	–	–	–	–	↓	–	↑	2
15	–	–	–	–	–	–	↑	–	↓	2
16	–	–	–	–	–	–	↓	↓	–	2
17	–	–	–	–	–	–	↑	↑	–	2
	E	E	E	E	E	E	A	A	A	

α_r	configuration 2									CI
1	–	–	–	–	–	–	–	–	–	0
2	–	–	–	–	–	↓	–	–	–	1
3	–	–	–	–	–	↑	–	–	–	1
4	–	–	–	–	–	↑↓	–	–	–	2
5	–	–	–	–	–	–	↓	–	–	1
6	–	–	–	–	–	–	↑	–	–	1
7	–	–	–	–	–	↓	↓	–	–	2
8	–	–	–	–	–	↑	↑	–	–	2
9	–	–	–	–	–	↑	↓	–	–	2
10	–	–	–	–	–	↓	↑	–	–	2
11	–	–	–	–	–	–	↑↓	–	–	2
12	–	–	–	–	–	–	–	–	↓	1
13	–	–	–	–	–	–	–	–	↑	1
14	–	–	–	–	–	↓	–	–	↓	2
15	–	–	–	–	–	↑	–	–	↑	2
16	–	–	–	–	–	↑	–	–	↓	2
17	–	–	–	–	–	↓	–	–	↑	2
	E	E	E	E	E	A	A	E	A	

α_r	configuration 3									CI
1	–	–	–	–	–	–	–	–	–	0
2	–	–	–	↓	–	–	–	–	–	1
3	–	–	–	↑	–	–	–	–	–	1
4	–	–	–	↑↓	–	–	–	–	–	2
5	–	–	–	–	↓	–	–	–	–	1
6	–	–	–	–	↑	–	–	–	–	1
7	–	–	–	↑	↓	–	–	–	–	2
8	–	–	–	↓	↑	–	–	–	–	2
9	–	–	–	–	↑↓	–	–	–	–	2
10	↓	–	–	–	–	–	–	–	–	1
11	↑	–	–	–	–	–	–	–	–	1
12	↓	–	–	↑	–	–	–	–	–	2
13	↑	–	–	↓	–	–	–	–	–	2
14	↓	–	–	–	↑	–	–	–	–	2
15	↑	–	–	–	↓	–	–	–	–	2
16	↑↓	–	–	–	–	–	–	–	–	2
	A	E	E	A	A	E	E	E	E	

Table 6: Three different subsets of states are formed from $M_r = 16$ states (or $M_r = 17$ due to spin reflection symmetry) expressed explicitely in an one-orbital basis which together with the $M_l q^2 = 64$ $(l)\bullet\bullet$ subsystem states fulfill the conservation of total number of particles with up and down spins. Labels E and A stands for empty and active orbitals respectively. In the very right columns the corresponding CI levels are indicated. The CAS vectors describing the three configurations are $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)$, $(2, 1, 9, 10, 3, 12, 11, 6, 5, 4, 8, 7)$, and $(7, 8, 4, 5, 6, 11, 12, 3, 10, 9, 1, 2)$, respectively.

of the system block, $S^{(\mathrm{L})}$, depends on the basis states used to construct the (R) environment block, as it increased from 10^{-5} to 0.22. Therefore, $S^{(\mathrm{L})}$ should be maximized by finding the best representation of the environment block for a given superblock configuration and target state, i.e, to get as close as possible to the exact solution, in the present case to $S^{(\mathrm{L})} = 0.28$.

This can be achieved by including highly entangled orbitals from the very beginning in the calculations. Therefore, in order to achieve fast and stable convergence the active space has to be expanded iteratively using orbitals with largest one-orbital entropy values. The sequence by which orbitals are taken into account is determined by the so called CAS-vector, which is simply a rendered sequence of orbital indices with decreasing one-orbital entropy value. The initial CAS vector can be determined based on the chemical character of the molecule or in a self-consistent fashion based on the single-orbital entropies. These features are incorporated in the DEAS procedure[127], see in section 4.3.3, starting with superblock configuration as shown in Fig. 36.

This approach has also been extended by including protocols based on the Configuration Interaction (CI) procedure[3,128]. In standard CI techniques, the trial wave function is written as a linear combination of determinants with expansion coefficients determined by requiring that the energy should be minimized[104]. The number of determinants included in the CI wave function expansion is increased systematically in order to achieve a better accuracy.

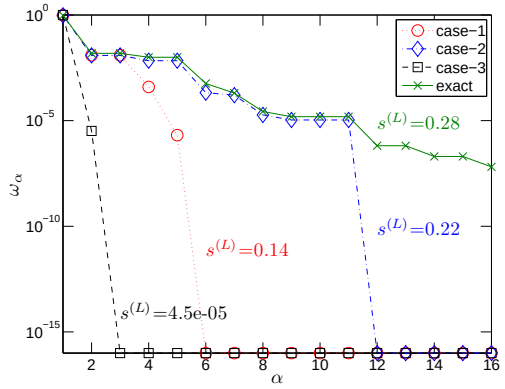


Figure 37: Eigenvalue spectrum of $\rho^{(\mathrm{L})}$ (and $\rho_{(\mathrm{R})}$) corresponding to the three subsets of environmnet states shown in table 6.

The exact wave function can be expressed as

$$\Psi_{\mathrm{CI}} = a_{\mathrm{HF}}\Phi_{\mathrm{HF}} + \sum_S a_S\Phi_S + \sum_D a_D\Phi_D + \sum_T a_T\Phi_T + \ldots \tag{196}$$

where determinants indicated by the subscripts S, D, T, Q are singly, doubly, triply, quadruply, etc. excited relative to the HF configuration.

If the HF orbitals are known, one can keep only those right block states which together with the lq^2 states of the $(l)\bullet\bullet$ composite system describe an excitation corresponding to a given CI-level. In the first iteration step this can be determined explicitly and the various CI excitations corresponding to basis states shown in Fig. 36 are given in the right column. In subsequent iteration steps, HF and non-HF orbitals can get mixed in renormalized multi-orbital basis states and they thus cannot be labeled by the CI excitation level. Nevertheless, the maximum CI level that block states could correspond to depends on the number of HF orbitals falling into the given block. Since the segment of the HF-orbitals belonging to the right(environment) block is known, the restricted subspace of the environment block can be formed for a given CI-level in the CI-DEAS procedure. Therefore, the right block contains states for a given CI-level while the total wave function can contain higher excitations as well due to the correlation between the two blocks. This procedure allows one to control the minimum CI-level to be used and a double optimization is carried out in each iteration step. On the one hand, the environment block states are constructed at each iteration step based on the left block basis states, thus they are optimized for the renormalized system (left) block. On the other hand, during the SVD step the left block states are optimized according to a well represented environment block, thus the reduced density matrix is well defined and block states can be selected efficiently based on the entropy considerations (DBSS, see in section 4.4.3). This procedure guarantees that several highly entangled orbitals are correlated from the very beginning and both static and dynamic correlations are taken into account, which helps to avoid convergence to local minima. Since a significant part of the correlation energy can be obtained in this way, usually at the end of the initialization procedure, i.e., after one-half sweep, chemical accuracy is reached. The starting value of M_r (M_{start}) is set prior to the calculation, but during the iteration procedure M_r is adjusted as $M_r = \max(M_l, M_{\mathrm{start}})$

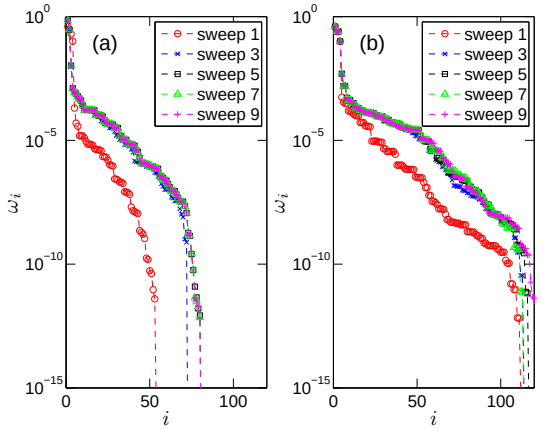


Figure 38: Eigenvalue spectrum of the reduced subsytem density matrix for a block of $l = 12$ contiguous orbitals as a function of DMRG sweeping for the LiF at $d_{\text{Li-F}} = 3.05$ a.u. using a non-optimized initialization procedure (a) and the CI-DEAS procedure (b).

in order to construct at least as many environment states as the left block has (to avoid zero Schmidt values).

The CI-DEAS procedure also has an important technical aspect. Based on the selected M_r basis states orbitals of the right block can be identified as doubly filled (D), empty (E) or active (A). If only the empty states appear in a given column of the configurational space as shown in table 6 the orbital is considered as empty, while if only the doubly filled state appears it is considered as doubly filled. Otherwise, the orbital is active. This is indicated explicitely in the last rows of table 6. It has been shown that empty orbitals can be neglected, while a partial summation over the doubly filled orbitals gives some corrections to the terms obtained by the partial summation over the active orbitals. Therefore, the effective size of the environment block can be reduced to the number of active orbitals[127,128]. Usually the number of active orbitals in the environment block range from 5 to 10 which allows one to use larger M_{start} without a significant increase in computational time.

As an example, the eigenvalue spectrum of the reduced subsytem density matrix for a block of $l = 12$ contiguous orbitals as a function of DMRG sweeping for the LiF CAS(6,25) at $d_{\text{Li-F}} = 3.05$ a.u. is shown in Fig. 38 using a non-optimized initialization procedure (a) and the CI-DEAS procedure (b).

Inclusion of the CI-DEAS procedure into the QC-TTNS method is straightforward. The only difference is that in a given iteration step of the wramup sweep instead of two, $z_i - 1$ environmnet blocks has to be formed.

4.4.8 Optimization of the sparsity using symmetries

As has been introduced in Sec. 4.1.4, symmetry operators (with eigenvalues Q, called quantum numbers) can be used to decompose the Hilbert space into subspaces (sectors)[52]. Therefore, the efficiency of the QC-DMRG and QC-TTNS methods can be increased significantly by applying quantum numbers. These include Abelian symmetries as particle number, spin projection[253], spin reflection[137] and Abelian point group symmetries[41,126,127] and even non-Abelian symmetries[123,152–155,192,205,217,220–223,230,231,245,245,262,264,276]. In the latter case, the sit-

uation, however, becomes more complicated.

If symmetry generators commute with each other, the eigenstates of the Hamiltonian form degenerate multiplets, $|\phi_{\alpha,\underline{Q}_\alpha,\underline{Q}^z_\alpha}\rangle$, that are classified by their label α, the quantum numbers $\underline{Q}_\alpha$, and the internal quantum number $\underline{Q}^z_\alpha$. The dimension of a subspace (sector) α depends uniquely on its quantum numbers $\underline{Q}_\alpha$, i.e. $\dim(\alpha) = \dim(\underline{Q}_\alpha)$. In the following, we use the shorthand notation introduced in Sec. 4.1 and write $|\phi_{\alpha,\underline{Q}_\alpha,\underline{Q}^z_\alpha}\rangle$, as $|\alpha,\underline{Q}_\alpha,\underline{Q}^z_\alpha\rangle$.

In general, the symmetry operators of the Hamiltonian are the representations $\mathcal{U}$ of the symmetry group $\mathcal{G}$ on the Hilbert space, acting as

$$\mathcal{U}(g)\,\mathbf{H}\,\mathcal{U}^{-1}(g) = \mathbf{H}, \tag{197}$$

where $\mathcal{U}(g)$ is the unitary representation for the symmetry $g \in \mathcal{G}$. Specially, if the symmetries are *local* in the sense that they decompose into unitary operators which commute with each other and act independently at different orbitals, then not only the whole Hamiltonian but also every local and interaction Hamiltonian are invariant under the group $\mathcal{G}$. Furthermore, $\mathcal{G}$ and correspondingly $\mathcal{U}$ can be decomposed into a direct product of Γ subgroups $\mathcal{G}_\gamma$ ($\gamma = 1,\ldots,\Gamma$), each acting independently on every orbital,

$$\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2 \times \cdots \times \mathcal{G}_\Gamma, \tag{198}$$

$$\mathcal{U}(g) = \prod_{\gamma=1}^{\Gamma} \mathcal{U}_\gamma(g_\gamma) = \prod_{\gamma=1}^{\Gamma}\prod_i \mathcal{U}_{\gamma,i}(g_\gamma). \tag{199}$$

Once a specific decomposition of the symmetry is obtained, Γ number of quantum numbers classify the irreducible subspaces (multiplets) of the subsystem Hamiltonians $\underline{Q} = \{Q^1, Q^2, \ldots, Q^\Gamma\}$ and states within the multiplet are then labeled by the *internal* quantum numbers $\underline{Q}^z = \{Q^{1,z}, Q^{2,z}, \ldots, Q^{\Gamma,z}\}$. The dimension of a subspace α depends uniquely on its quantum numbers $\underline{Q}_\alpha$, i.e. $\dim(\alpha) = \dim(\underline{Q}_\alpha) = \prod_{\gamma=1}^{\Gamma} \dim(Q^\gamma_\alpha)$.

Operators can also be arranged into irreducible tensor operators, and an irreducible tensor operator multiplet A is correspondingly described by quantum numbers $\underline{a}$, while members of the multiplet are labeled by $\underline{a}^z$ with $\underline{a}$ and $\underline{a}^z$ being Γ-component vectors. The Wigner–Eckart theorem[247,254–256] tells us that, apart from trivial group theoretical factors (Clebsch-Gordan coefficients), the matrix elements of the members of a given operator multiplet and states within two multiplets, $|\alpha, Q_\alpha, Q^z_\alpha\rangle$ and two multiplets, $|\alpha', Q_{\alpha'}, Q^z_{\alpha'}\rangle$ are simply related by

$$\left\langle \alpha, \underline{Q}_\alpha \underline{Q}^z_\alpha \right| A_{\underline{a},\underline{a}_z} \left| \alpha', \underline{Q}_{\alpha'} \underline{Q}^z_{\alpha'} \right\rangle = \langle \alpha \parallel A \parallel \alpha' \rangle \left\langle \underline{Q}_\alpha \underline{Q}^z_{\alpha'} \right| \underline{a}, \underline{a}_z; \underline{Q}_\alpha \underline{Q}^z_{\alpha'} \right\rangle \tag{200}$$

where $\langle \alpha \parallel A \parallel \alpha' \rangle$ denotes the reduced (invariant) matrix element of A, and the generalized Clebsch–Gordan coefficients are simply defined as

$$\left\langle \underline{Q}_\alpha\ \underline{Q}^z_\alpha \ \middle|\underline{a}\ \underline{a}^z; \underline{Q}_{\alpha'}\ \underline{Q}^z_{\alpha'} \right\rangle \equiv \prod_{\gamma=1}^{\Gamma} \langle Q^\gamma_\alpha\ Q^{\gamma,z}_\alpha\ \ |a^\gamma\ a^{\gamma,z}\ Q^\gamma_{\alpha'}; Q^{\gamma,z}_{\alpha'}\rangle\ . \tag{201}$$

In the presence of symmetries, one has to use the Clebsch-Gordan coefficients to build (L) block states from the block (l) and orbital $\bullet$ states that transform as irreducible multiplets

under the symmetry transformations, $\mathcal{U}(g)$,

$$\left|m_{(\mathrm{L})}, \underline{Q}_{m_{(\mathrm{L})}}, \underline{Q}^z_{m_{(\mathrm{L})}}\right\rangle \equiv \sum_{\underline{Q}^z_{m_{(\mathrm{l})}}, \underline{Q}^z_{\alpha_{l+1}}} \left\langle \underline{Q}_{m_{(\mathrm{L})}} \underline{Q}^z_{m_{(\mathrm{L})}} \middle| \underline{Q}_{\alpha_{l+1}} \underline{Q}^z_{\alpha_{l+1}}; \underline{Q}_{m_{(\mathrm{l})}} \underline{Q}^z_{m_{(\mathrm{l})}} \right\rangle^* \left|\alpha_{l+1}, \underline{Q}_{\alpha_{l+1}}, \underline{Q}^z_{\alpha_{l+1}}\right\rangle \otimes \left|m_{(\mathrm{l})}, \underline{Q}_{m_{(\mathrm{l})}}, \underline{Q}^z_{m_{(\mathrm{l})}}\right\rangle, \tag{202}$$

Therefore, subsystem Hamiltonians have a block-diagonal structure and subsystem reduced density matrices are also scalar under symmetry operators. This decomposition property is crucial for using symmetries in the QC-DMRG and QC-TTNS calculations in order to boost their performance.

To give a simple example, let us take into account the spin and charge symmetries, i.e., $\mathcal{G} = \mathcal{G}_{\mathrm{spin}} \times \mathcal{G}_{\mathrm{charge}}$. If we use only $\mathcal{G}_{\mathrm{spin}} = U(1)$ and $\mathcal{G}_{\mathrm{charge}} = U(1)$ symmetries then one has two hopping operators, $\mathbf{c}^\dagger_{i,\uparrow}$ and $\mathbf{c}^\dagger_{i,\downarrow}$ as defined in Sec. 4.1. In contrast to this, if we use $\mathcal{G}_{\mathrm{spin}} = SU(2)$ spin symmetry, (while $\mathcal{G}_{\mathrm{charge}} = U(1)$ remains the same as before) $\mathcal{U} = \mathcal{U}_{\mathrm{spin}}\mathcal{U}_{\mathrm{charge}}$, with $\mathcal{U}_{\mathrm{spin}} = \prod_i \mathcal{U}_{\mathrm{spin},i}$, then only one hopping operator remains since matrix elements of $\mathbf{c}^\dagger_{i,\uparrow}$ and $\mathbf{c}^\dagger_{i,\downarrow}$ are related with each other by symmetry and they form a single operator multiplet $\mathbf{c}^\dagger_i = \{\mathbf{c}^\dagger_{i,\downarrow}, \mathbf{c}^\dagger_{i,\uparrow}\}$ of spin 1/2. The matrix elements of such multiplet are determined using the Wigner-Eckart theorem

$$\langle \mu \parallel \mathbf{c}^\dagger_i \parallel \nu \rangle = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -\sqrt{2} & 0 \end{pmatrix}, \tag{203}$$

where the original $\mathbb{C}^4$ space is reduced to $\mathbb{C}^3$ since only three basis states remain $\mu, \nu \in \{|-\rangle, |\uparrow\rangle, |\uparrow\downarrow\rangle\}$. When the system is half-filled, utilization of $\mathcal{G}_{\mathrm{charge}} = SU(2)$ symmetriy besides the $\mathcal{G}_{\mathrm{spin}} = SU(2)$ spin symmetry is straightforward. In this case the hopping operator becomes a 2×2 matrix

$$\begin{pmatrix} 0 & \sqrt{2} \\ -\sqrt{2} & 0 \end{pmatrix}, \tag{204}$$

where the original $\mathbb{C}^4$ space further reduces to $\mathbb{C}^2$ since only two basis states remain, $|\phi_\mu\rangle, |\phi_\nu\rangle \in \{|\uparrow\rangle, |\uparrow\downarrow\rangle\}$. A detailed derivation of reduced operators and construction of the block states prepared as a pedagogical introdcution to the field can be found in the literature[123], and the related free `C++` sourcecode can be downloaded from `http://www.phy.bme.hu/~dmnrg/`[134]. Other free source codes with SU(2) spin symmetries are also available[216,260]. Utilization of symmetries allows one to target states with given symmetries and to keep M number of multiplets which correspons to significantly more $U(1)$ states what is crucial in order to achieve good numerical accuracy.

4.4.9 Stability of the wavefunction

Traditional post-HF quantum chemical methods like CI or Couple Clusters (CC) systematically improve a refernce wavefunction (often only the HF determinant, as in (196)) by inclusion of single, double, and higher excitations in the wave operator. In case of CI the wave operator takes a linear form, while CC uses a more sophisticated exponential ansatz.

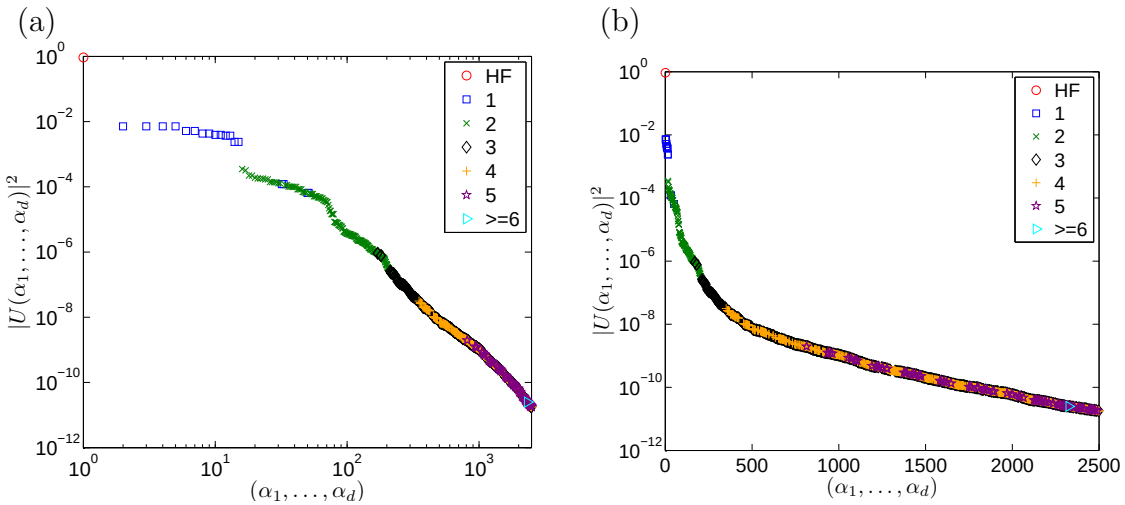


Figure 39: The $|U(\alpha_1,\ldots\alpha_d)|^2$ weight of the d-orbital basis states corresponding to the various CI-excitations are shown by different colors in a descending order in a log-log (a) and in a log-lin scale (b) for the LiF for $d_{\text{Li-F}} = 3.05$ with CAS(6,12). HF state, SCI, DCI, etc are indicated by red, blue, green, etc. colors respectively.

In contrast to these, QC-DMRG and QC-TTNS take into account all the various excitations picking up the most important ones by minimizing the energy. As an example, the $|U(\alpha_1,\ldots\alpha_d)|^2$ weights casted according to an excitation level with respect to the HF refernce wavefunction are shown in Fig. 39. It demonstrates that higher excitation levels can be important to provide a qualitatively correct description of the wavefunction. (The elements of the full tensor $U(\alpha_1,\ldots\alpha_d)$ can be extracted, according to Eq. (176). But note that recovering all components of $U(\alpha_1,\ldots\alpha_d)$ cannot be done efficiently as its size scales exponentially. However as a good approximation of the full CI wavefunction, the Monte Carlo algorithm was used to recover the most important tensor components[27].)

As a consequence, if the accuracy threshold of the calculation is lowered, the structure of the wave function is retained in essence. Since the DBSS procedure takes care of the change in the entanglement as the system parameters are adjusted, for example, when the bond length in LiF is changed, the various calculated quantities are continuous functions for a given $\delta\varepsilon_{\text{TR}}$. For the ionic–neutral curve crossing in LiF[126] this has been demonstrated for the two lowest $^1\Sigma^+$ states and the dipole moment function as illustrated in Fig. 40(a) and (b) for $\delta\varepsilon_{\text{TR}} = 10^{-6}$ and $M_{\text{min}} = 64$. In addition, when parameters were cutted drastically and very small value of M_{min} and large $\delta\varepsilon_{\text{TR}}$ were used the dipole moment deviated more significantly from the full-CI results but they remained continuous even close to the avoided crossing. Therefore, the most important components of the wave function are included by the SVD procedure which provides a stable representation of the wavefunction. Similar results have been reported for the QC-TTNS method[170].

4.4.10 Possible black-box QC-DMRG and QC-TTNS

A possible black-box QC-DMRG and QC-TTNS can be composed of two phases: the *preprocessing phase* in which the ordering, network topology and CAS-vector are optimized using

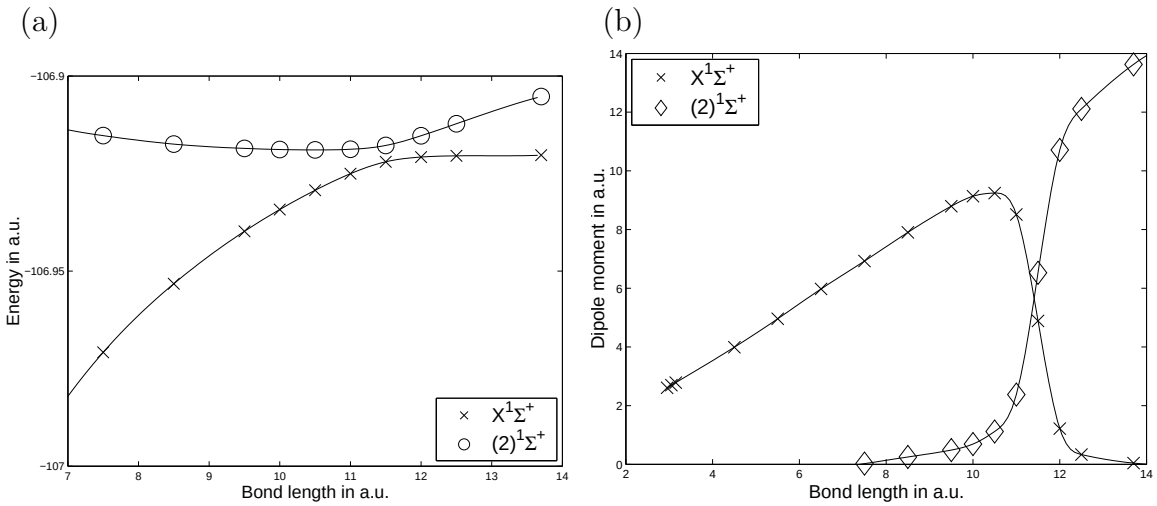


Figure 40: (a) Energy of the two lowest $^1\Sigma^+$ states as a function of the bond length obtained with $\delta\varepsilon_{\rm TR} = 10^{-6}$ and $M_{\rm min} = 16$ using the DBSS procedure. (b) The corresponding dipole moment functions.

fixed small number of block states and the *production phase* in which an accurate calculation is performed using the DBSS procedure in order to reach an a priory set error margin. In the preprocessing phase, one can use the ordering for which the integral files were generated and a random CAS vector using limited number of block states. After a full sweep the one-orbital entropy can be calculated from which the CAS vector can be determined. In a similar way the two-orbital mutual information and the optimal ordering can be calculated using the Fiedler vector. Next a DMRG calculation can be carried out with the optimized ordering and CAS-vector and the whole cycle is repeated until we obtain lower total energy. In the next step this procedure is repeated, but with larger number of block states. The preprocessing phase takes only a small fraction of the total computational time.

4.5 Miscellaneous

4.5.1 Simulation of real materials, geometrical optimization and excited states

As an example, we demonstrate on *poly-diacetylene* (PDA) chains that MPS based methods can be used very efficiently to simulate strongly anisotropic materials in terms of effective Hamiltonians.

PDA chains dispersed with low concentration in their monomer single-crystal matrix are prototypical quasi one-dimensional materials[25,33,207,213]. The structural disorder in the chains and their surrounding matrix is tiny, thus these materials form the perfect testing-ground for theoretical model studies describing interacting electrons on perfectly ordered chains. In addition, the electronic excitation energies of the diacetylene monomers are much higher than those of the polymer, and the electronic excitations of the chain in the energy range of visible light can be measured with a very high accuracy[224]. Polymerization induced by temperature or ultraviolet light is shown in Fig. 41.

The opto-electronic properties of the PDAs are determined by two main correlation effects: the mutual interaction of the electrons and their interaction with the lattice poten-

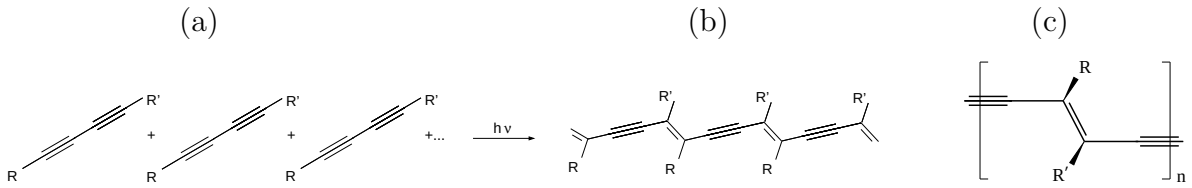


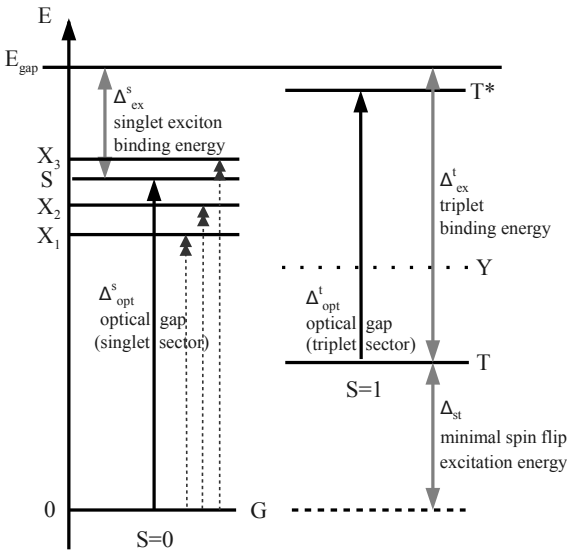
Figure 41: (a) Single crystals of diacetylene monomers prepared experimentally where the monomer unit has the structure R-C≡C-C≡C-R. In the *n*BCMU family R = $(CH_2)_n$-OCONH-CH_2-COO-$(CH_2)_3CH_3$. (b) Polymerization induced by temperature or ultraviolet light leading to polydiacetylenes $(C_4R_2)_x$. (c) Lewis structure of a poly-diacetylene unit cell with single, double and triple bond lengths $r_t = 1.20$Å, $r_d = 1.36$Å and $r_s = 1.43$Å, respectively.

tial[60,213]. In contrast to inorganic semiconductors, the exciton binding energy in PDAs amounts to about 20% of the single-particle gap, thus Coulomb interaction is substantial and effective and the electron-electron interaction must be treated very accurately. Due to such high computational demand earlier attempts based on density-functional theory calculations of the bare band structure in local-density approximation (LDA) failed to reproduce the experimentally measured excitation spectrum[202,235].

In contrast to this, using the DMRG method and by correlating some 100 electrons on 100 orbitals together with a geometrical optimization based on the Hellmann–Feynman theorem, i.e., by minimizing the force-field induced by the electron distribution[196], very accurate energy spectrum can be obtained[18]. This is shown in Fig. 42 and the experimentally measured and DMRG calculated results are also summarized in the corresponding table. In addition, the calculated geometrical structure agrees perfectly with the experimental data, i.e., the single, double and triple bonds are estimated as $r_t = 1.22$, $r_d = 1.37$ and $r_s = 1.43$.

4.5.2 Four-component density matrix renormalization group

Quite recently, the first implementation of the *relativistic* quantum chemical two- and four-component density matrix renormalization group algorithm (2c- and 4c-DMRG) has also been presented[109]. This method includes a variational description of scalar-relativistic effects and spin–orbit coupling. By correlating 14 electrons on 94 spinors and employing the Dirac–Coulomb Hamiltonian with triple-ζ quality basis, the *Potential Energy Surface* (PES) and spectroscopic constants have been obtained for the *thallium hydride molecule*. Utilizing the various entanglement based optimization techniques discussed in Sec. 4.4, the CCSD reference energy has been reproduced even after the first DMRG sweep as is shown in Fig. 43. Although the 4c-CCSDTQ reference energy could not be reached with a maximum of $M = 4500$ block states, the resulting 4c-DMRG potential energy curve did not only effectively reproduced the shape of the 4c-CCSDTQ potential energy curve but also yielded accurate spectroscopic constants as extracted from a fourth-order polynomial fit. Since QC-DMRG picks up all excitations required to describe the wave function to a given accuracy the general structure of the wave function is preserved and could have been determined even with smaller M values. By making the best of entanglement optimization the new 2c- and 4c-DMRG method is expected to become an efficient approach for heavy-element molecules that exhibit rather strong multi-configurational character in their ground- and excited states.



Energy/eV	3BCMU	DMRG
E_{X_1}	*1.5*	1.74 [1.94]
E_{X_2}	*1.7*	1.85[a] [2.0][a]
$E_{\mathrm{S}} = \Delta^{\mathrm{s}}_{\mathrm{opt}}$	**1.896**	2.00 [2.05]
E_{X_3}	*2.0*	
E_{gap}	**2.482**	2.45 [2.47]
$\Delta^{\mathrm{s}}_{\mathrm{ex}} = E_{\mathrm{gap}} - E_{\mathrm{S}}$	**0.586**	0.45 [0.42]
$E_{\mathrm{T}} = \Delta_{\mathrm{st}}$	*1.0* $\pm$ *0.05*	1.00 [1.06]
$E_{\mathrm{T}^*} = \Delta_{\mathrm{st}} + \Delta^{\mathrm{t}}_{\mathrm{opt}}$	*2.36* $\pm$ *0.05*	2.25
$\Delta^{\mathrm{t}}_{\mathrm{opt}}$	**1.360**	1.25 [1.28]
$\Delta^{\mathrm{t}}_{\mathrm{ex}} = E_{\mathrm{gap}} - E_{\mathrm{T}}$	*1.5* $\pm$ *0.05*	1.45 [1.40]

Figure 42: (left) Energy levels of in-gap states in the spin-singlet and spin-triplet sectors. Single-tip arrows: optical absorption spectroscopy; double-tip arrows: two-photon absorption spectroscopy. Double arrows: binding energies (gaps). G: singlet ground state (1^1A_g); S: singlet exciton (1^1B_u); X_1, X_2, X_3: singlet dark states (m^1A_g); T: triplet ground state (1^3B_u); T^*: optical excitation of the triplet ground state (1^3A_g); Y: dark triplet state (m^3B_u). (right) Excitation energies in 3BCMU at low temperatures. All energies are measured in eV relative to the energy of the ground state, $E_{\mathrm{G}} = 0$. Bold number: directly measured; italic number: estimate. For DMRG results the numbers in square brackets give the excitation energy for the rigid-lattice transition from G (E_{gap}, E_{S}, $E_{\mathrm{X}_{1,2}}$, E_{T}) and from T ($\Delta^{\mathrm{t}}_{\mathrm{opt}}$).

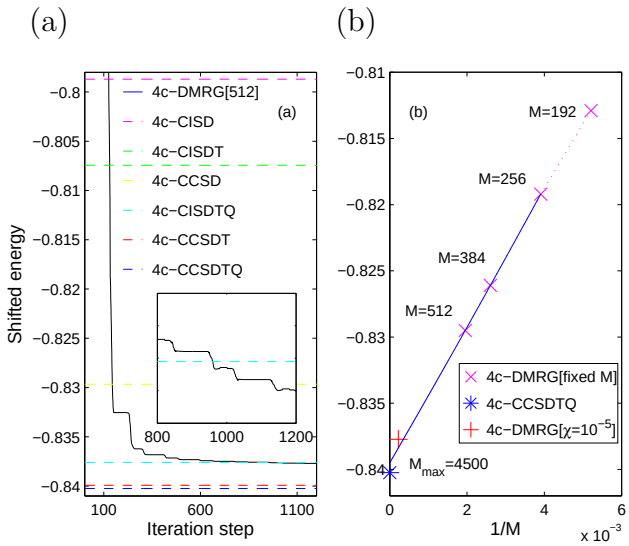


Figure 43: (a) Convergence of the ground state energy (shifted by 20275 E_h) as a function of iteration steps of the 4c-DMRG(14,94) ($M_\mathrm{max} = 4500$, $M_\mathrm{min} = 1024$, $M_\mathrm{min}^\mathrm{DEAS} = 2048$, $\chi = 10^{-5}$) approach at $r_e^\mathrm{exp} = 1.872$ Å. Reference energies calculated by various CI and CC wave function models are also given as horizontal lines. The inset shows that the 4c-DMRG energy drops below the 4c-CI-SDTQ energy. (b) Extrapolation of DMRG energies E(M)-20275 E_h for fixed M values towards the limit $E(M \to \infty)$-20275 E_h. Figure is taken from arxiv:1312.0970.

Development of a 2c- and 4c-TTNS method is straightforward.

4.5.3 Possible technical developments: hybrid CPU/GPU parallelization

The original DMRG algorithm, introduced by S. R. White, was formulated as a single threaded algorithm[249]. In the past various works have been carried out to accelerate the DMRG algorithm on shared[10,90] and distributed memory[35,114,200,267] architectures. One of the first parallelizations was converting the projection operation to matrix-matrix multiplications and accelerating them via OpenMP interface[90]. A similar approach has been presented for distributed memory environment (up-to 1024 cores) optimizing the communication between the cores[267], while the acceleration of the computation of correlation function has also been investigated[200]. A novel direction for parallelization via a modification of the original serial DMRG algorithm have also been introduced[225].

Graphical Processing Unit (GPU) has been successfully employed in neighboring research areas to accelerate matrix operations. GPU is used to accelerate tensor contractions in *Plaquette Renormalization States* (PRS)[273], which can be regarded as an alternative technique to tensor network states (TNS) or the DMRG algorithm. The *second-order Spectral Projection* (SP2) algorithm has been accelerated, which is an alternative technique to calculate the density matrix via a recursive series of generalized matrix-matrix multiplications[34]

Quite recently, it has been investigated how the DMRG method can utilize the enormous computing capabilities of novel kilo-processor architectures: *Graphical Processing Unit* (GPU) and *Field-Programmable Gate Array* (FPGA)[175]. In case of GPU a smart hybrid

CPU-GPU acceleration has been presented, which tolerates problems exceeding the GPU memory size, consequently, supporting wide range of problems and GPU configurations. Contrary to earlier acceleration attempts not only the projection operation was accelerated, but further parts of the diagonalization were also computed on the GPU. Reported results on the one-dimensional Hubbard model for a mid-range (Intel Core-i7 2600 3.4 GHz CPU + NVidia GTX 570 GPU) and on a high-end configuration (Intel Xeon E5-2640 2.5 GHz CPU + NVidia K20 GPU) showed that if the workload is properly distributed (see Fig. 44) the mid-range configuration with GPU can be approximately 2.3-2.4 times faster than without GPU, while the high-end configuration can be accelerated by 3.4-3.5 times using the GPU.

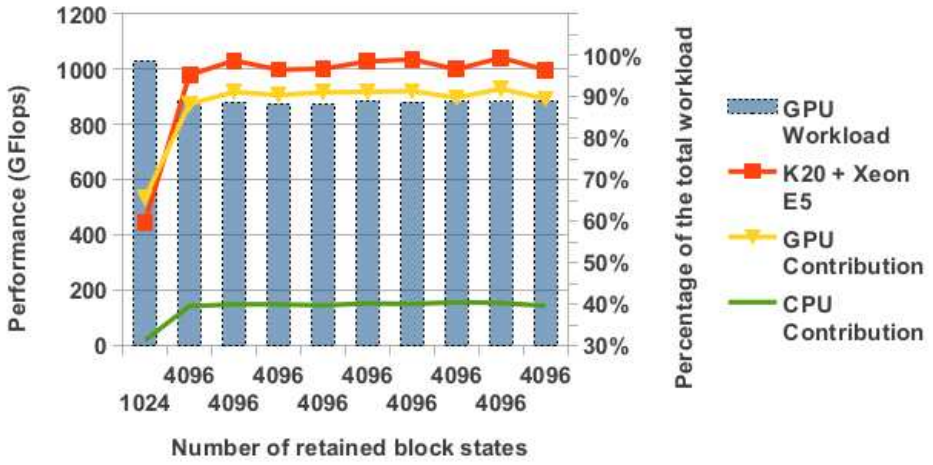


Figure 44: Performance results of the hybrid CPU-GPU acceleration of the projection operation for the Hubbard model on Intel Xeon E5-2640 2.5GH CPU + NVidia K20 GPU: 1071 GFlops and $\times 3.5$ speedup is reached. (Theoretical maximum is 1.17 TFlops) Blue bars associated to the secondary vertical axis indicate the ratio of the current GPU workload.

The GPU architecture has been found to be a promising accelerator, as the most time-dominant step of the algorithm, the projection operation, can be formulated as independent dense matrix multiplications, which are ideal workload for GPUs. Moreover, in case of high-end GPUs the acceleration of the projection is so remarkable, that it is worth to consider the acceleration of the rest of the algorithm to obtain a decent overall speed-up. Therefore, extensions to treat ab-initio quantum chemical applications and a straightforward generalization of the algorithm to accelerate tensor network state algorithms [180] are promising research directions.

5 Summary and outlook

In the past decade, we have witnessed a breakthrough in electronic structure calculations due to the *Density Matrix Renormalization Group* (DMRG) method which has become a viable alternative to conventional multiconfiguration wave function approaches. Inclusion

of the concepts of entanglement from quantum information theory has paved the road for identifying highly correlated molecular orbitals leading to an efficient construction of active spaces and for characterizing the various types of correlation effects relevant for chemical bonding. Quite recently, a reformulation of DMRG in terms of *Matrix Product States* (MPS) has shown that it is only one special case in a much more general set of methods, the *Tensor Network States* (TNS), which is expected to even outperform DMRG/MPS in the near future.

A special class of such ansatz states are the *Tree Tensor Network States* (TTNS). The mathematically rigorous analysis of these tensor trees has been completed only partially and many open questions remain, concerning for example numerical procedures, but also more theoretical concepts of differential and algebraic geometry.

In the *quantum chemsitry version* of the method (QC-TTNS), the wave function with variable tensor order is formulated as products of tensors in a multiparticle basis spanning a truncated Hilbert space of the original CAS-CI problem. The tree-structure is advantageous since the distance between two arbitrary orbitals in the tree scales only logarithmically with the number of orbitals, whereas the scaling is linear in the MPS array. Therefore, the TTNS ansatz is better suited for multireference problems with numerous highly correlated orbitals.

The underlying benefits of QC-TTNS is, however, far from fully exploited and the optimization of the method is far more complicated. Due to the more advanced topology, several optimization tasks and problems arise which do not have counterparts in the MPS formulation. Therefore, there is a tedious work still ahead of us.

6 Acknowledgments

This research was supported by the European Research Area Chemistry (ERA-Chemistry) in part by the Hungarian Research Fund (OTKA) under Grant No. NN110360 and K100908, the DFG SCHN 530/9-1 project under Grant No. 10041620 and FWF-E1243-N19. V.M. and F.V. acknowledge support from the SFB-ViCoM project. Ö.L. acknowledges support from the Alexander von Humboldt foundation and from ETH Zürich during his time as a visiting professor. The authors are also grateful for Libor Veis, Jenő Sólyom, Péter Nagy, and Tibor Szilvási for helpful discussions and comments.

References

[1] International workshop on recent progress and prospects in density-matrix renormalization, 2004. 2 - 13 August 2004, Lorentz Center, Leiden University, The Netherlands.

[2] International workshop on density matrix renormalization group and other advances in numerical renormalization group methods, 2010. 23 August - 3 September 2010, Beijing, China.

[3] Workshop on tensor network methods for quantum chemistry, 2010. 29 - 31 March 2010, CECAM-ETHZ, Zurich, Switzerland.

[4] Workshop on entanglement based approaches in quantum chemistry (ebaqc), 2012. 3 - 6 September 2012, MPIPKS-Dresden, Germany.

[5] Workshop on tensor network algorithms in computational physics and numerical analysis, 2013. 15 - 17 May 2013, CECAM-ETHZ, Zurich, Switzerland.

[6] Workshop on theoretical chemistry: New wavefunction methods and entanglement optimizations in quantum chemistry, 2014. 18 - 21 February 2014, Mariapfarr, Austria.

[7] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008.

[8] Ian Affleck, Tom Kennedy, Elliott H. Lieb, and Hal Tasaki. Rigorous results on valence-bond ground states in antiferromagnets. *Phys. Rev. Lett.*, 59:799–802, Aug 1987.

[9] D. R. Alcoba, R. C. Bochicchio, L. Lain, and A. Torre. On the measure of electron correlation and entanglement in quantum chemistry based on the cumulant of the second-order reduced density matrix. *The Journal of Chemical Physics*, 133(14):144104, 2010.

[10] Gonzalo Alvarez. Implementation of the su(2) hamiltonian symmetry for the dmrg algorithm. *Computer Physics Communications*, 183(10):2226 – 2232, 2012.

[11] Luigi Amico, Rosario Fazio, Andreas Osterloh, and Vlatko Vedral. Entanglement in many-body systems. *Rev. Mod. Phys.*, 80:517–576, May 2008.

[12] Frithjof B. Anders and Avraham Schiller. Real-time dynamics in quantum-impurity systems: A time-dependent numerical renormalization-group approach. *Phys. Rev. Lett.*, 95:196801, Oct 2005.

[13] Andreas Arnold and Tobias Jahnke. On the approximation of high-dimensional differential equations in the hierarchical tucker format. *BIT Numerical Mathematics*, 54(2):305–341, 2014.

[14] J. Atkins, E. Boman, and B. Hendrickson. A spectral algorithm for seriation and the consecutive ones problem. *SIAM Journal on Computing*, 28(1):297–310, 1998.

[15] G. Barcza, Ö. Legeza, K. H. Marti, and M. Reiher. Quantum-information analysis of electronic states of different molecular structures. *Phys. Rev. A*, 83:012508, Jan 2011.

[16] G. Barcza, Ö. Legeza, R. M. Noack, and J. Sólyom. Dimerized phase in the cross-coupled antiferromagnetic spin ladder. *Phys. Rev. B*, 86:075133, Aug 2012.

[17] G. Barcza, R. M. Noack, J. Sólyom, and Ö. Legeza. Entanglement patterns and generalized correlation functions in quantum many body systems. *arXiv [cond-mat.str-el]*, page 1406.6643, 2014.

[18] Gergely Barcza, William Barford, Florian Gebhard, and Örs Legeza. Excited states in polydiacetylene chains: A density matrix renormalization group study. *Phys. Rev. B*, 87:245116, Jun 2013.

[19] Charles W. Bauschlicher and Stephen R. Langhoff. Full configurationinteraction study of the ionicneutral curve crossing in LiF. *The Journal of Chemical Physics*, 89(7):4246–4254, 1988.

[20] M. H. Beck, A. Jäckle, G. A. Worth, and H. d. Meyer. The multiconfiguration time-dependent Hartree (MCTDH) method: A highly efficient algorithm for propagating wavepackets. *Rep*, 324:1–105, 1999.

[21] John S. Bell. On the einstein podolsky rosen paradox. *Physics*, 1:195, 1967.

[22] Udo Benedikt, Alexander Auer, Mike Espig, and Wolfgang Hackbusch. Tensor decomposition in post-Hartree Fock methods. Pt. 1 : Two-electron integrals and MP2. *The journal of chemical physics*, 134(5):054118, 2011.

[23] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63:012307, Dec 2000.

[24] O. Berkman and U. Vishkin. Recursive star-tree parallel data structure. *SIAM Journal on Computing*, 22(2):221–242, 1993.

[25] David Bloor and R.R. Chance, editors. *Polydiacetylenes: Synthesis, Structure and Electronic Properties*, volume 102 of *Nato Science Series E*. Nijhoff, Dordrecht, 1985.

[26] Katharina Boguslawski, Konrad H. Marti, Örs Legeza, and Markus Reiher. Accurate ab initio spin densities. *Journal of Chemical Theory and Computation*, 8(6):1970–1982, 2012. PMID: 22707921.

[27] Katharina Boguslawski, Konrad H. Marti, and Markus Reiher. Construction of casci-type wave functions for very large active spaces. *The Journal of Chemical Physics*, 134(22):224101, 2011.

[28] Katharina Boguslawski and Paweł Tecmer. Orbital entanglement in quantum chemistry. *arXiv [physics.chem-ph]]*, page 1409.8017, 2014.

[29] Katharina Boguslawski, PawełTecmer, Gergely Barcza, Örs Legeza, and Markus Reiher. Orbital entanglement in bond-formation processes. *Journal of Chemical Theory and Computation*, 9(7):2959–2973, 2013.

[30] Katharina Boguslawski, Pawel Tecmer, Örs Legeza, and Markus Reiher. Entanglement measures for single- and multireference correlation effects. *The Journal of Physical Chemistry Letters*, 3(21):3129–3135, 2012.

[31] S. F. Boys. Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.*, 32:296–299, Apr 1960.

[32] Pasquale Calabrese and John Cardy. Entanglement entropy and quantum field theory. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(06):P06002, 2004.

[33] Hans-Joachim Cantow, editor. *Polydiacetylenes*, volume 63 of *Advances in Polymer Science*. Springer Berlin Heidelberg, 1984.

[34] M. J. Cawkwell, E. J. Sanville, S. M. Mniszewski, and Anders M. N. Niklasson. Computing the density matrix in electronic structure theory on graphics processing units. *Journal of Chemical Theory and Computation*, 8(11):4094–4101, 2012.

[35] Garnet Kin-Lic Chan. An algorithm for large scale density matrix renormalization group calculations. *The Journal of Chemical Physics*, 120(7):3172–3178, 2004.

[36] Garnet Kin-Lic Chan. Density matrix renormalisation group lagrangians. *Phys. Chem. Chem. Phys.*, 10:3454–3459, 2008.

[37] Garnet Kin-Lic Chan. Low entanglement wavefunctions. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(6):907–920, 2012.

[38] Garnet Kin-Lic Chan, Jonathan J. Dorando, Debashree Ghosh, Johannes Hachmann, Eric Neuscamman, Haitao Wang, and Takeshi Yanai. An introduction to the density matrix renormalization group ansatz in quantum chemistry. In Stephen Wilson, Peter J. Grout, Jean Maruani, Gerardo Delgado-Barrio, and Piotr Piecuch, editors, *Frontiers in Quantum Systems in Chemistry and Physics*, volume 18 of *Progress in Theoretical Chemistry and Physics*. Springer, 2008.

[39] Garnet Kin-Lic Chan and Martin Head-Gordon. Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group. *The Journal of Chemical Physics*, 116(11):4462–4476, 2002.

[40] Garnet Kin-Lic Chan and Martin Head-Gordon. Exact solution (within a triple-zeta, double polarization basis set) of the electronic schrödinger equation for water. *The Journal of Chemical Physics*, 118(19):8551–8554, 2003.

[41] Garnet Kin-Lic Chan, Mihály Kállay, and Jürgen Gauss. State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve. *The Journal of Chemical Physics*, 121(13):6110–6116, 2004.

[42] Garnet Kin-Lic Chan and Sandeep Sharma. The density matrix renormalization group in quantum chemistry. *Annual Review of Physical Chemistry*, 62(1):465–481, 2011. PMID: 21219144.

[43] Garnet Kin-Lic Chan and Troy Van Voorhis. Density-matrix renormalization-group algorithms with nonorthogonal orbitals and non-hermitian operators, and applications to polyenes. *The Journal of Chemical Physics*, 122(20):204101, 2005.

[44] Garnet Kin-Lic Chan and Dominika Zgid. Chapter 7 the density matrix renormalization group in quantum chemistry. volume 5 of *Annual Reports in Computational Chemistry*, pages 149 – 162. Elsevier, 2009.

[45] Hitesh J. Changlani, Jesse M. Kinder, C. J. Umrigar, and Garnet Kin-Lic Chan. Approximating strongly correlated wave functions with correlator product states. *Phys. Rev. B*, 80:245116, Dec 2009.

[46] Eric Chitambar, Debbie Leung, Laura Maninska, Maris Ozols, and Andreas Winter. Everything you always wanted to know about locc (but were afraid to ask). *Communications in Mathematical Physics*, 328(1):303–326, 2014.

[47] P.G. Ciarlet and J.L. Lions. *Handbook of Numerical Analysis: Computational chemistry*. Computational Chemistry: Reviews of Current Trends. North-Holland, 2003.

[48] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969.

[49] Dajana Conte and Christian Lubich. An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44:759–780, 7 2010.

[50] Philippe Corboz, Glen Evenbly, Frank Verstraete, and Guifré Vidal. Simulation of interacting fermions with entanglement renormalization. *Phys. Rev. A*, 81:010303, Jan 2010.

[51] Philippe Corboz and Guifré Vidal. Fermionic multiscale entanglement renormalization ansatz. *Phys. Rev. B*, 80:165129, Oct 2009.

[52] J. F. Cornwell. *Group Theory in Physics, An Introduction*. Academic Press, 1997.

[53] Christopher J. Cramer, Marta Woch, Piotr Piecuch, Cristina Puzzarini, and Laura Gagliardi. Theoretical models on the cu2o2 torture track: mechanistic implications for oxytyrosinase and small-molecule analogues. *The Journal of Physical Chemistry A*, 110(5):1991–2004, 2006. PMID: 16451035.

[54] C. Da Silva and F.J. Hermann. Optimization on the hierarchical tucker manifold - applications to tensor completion. *arXiv [math.NA]*, page 1405.2096, 2014.

[55] Ernest R. Davidson. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *Journal of Computational Physics*, 17(1):87 – 94, 1975.

[56] S. V. Dolgov, Boris N. Khoromskij, and Ivan V. Oseledets. Fast solution of parabolic problems in the tensor train/quantized tensor train format with initial application to the fokker-planck equation. *SIAM J. Scientific Computing*, 34(6), 2012.

[57] Jonathan J. Dorando, Johannes Hachmann, and Garnet Kin-Lic Chan. Targeted excited state algorithms. *The Journal of Chemical Physics*, 127(8):084109, 2007.

[58] Jonathan J. Dorando, Johannes Hachmann, and Garnet Kin-Lic Chan. Analytic response theory for the density matrix renormalization group. *The Journal of Chemical Physics*, 130(18):184111, 2009.

[59] Sidney D. Drell, Marvin Weinstein, and Shimon Yankielowicz. Quantum field theories on a lattice: Variational methods for arbitrary coupling strengths and the ising model in a transverse magnetic field. *Phys. Rev. D*, 16:1769–1781, Sep 1977.

[60] Francois Dubin, Romain Melet, Thierry Barisien, Roger Grousson, Laurent Legrand, Michel Schott, and Valia Voliotis. Macroscopic coherence of a single exciton state in an organic quantum wire. *Nat Phys*, 2:32–35, 2006.

[61] Jorge Dukelsky and Stuart Pittel. The density matrix renormalization group for finite fermi systems. *Reports on Progress in Physics*, 67(4):513, 2004.

[62] Corinne Duperrouzel, Paweł Tecmer, Katharina Boguslawski, Gergerly Barcza, Örs Legeza, and Paul W. Ayers. A quantum informational approach for dissecting chemical reactions. *arXiv [physics.chem-ph]*, page 1409.4867, 2014.

[63] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777, 1935.

[64] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82:277–306, Feb 2010.

[65] Mike Espig, Wolfgang Hackbusch, Stefan Handschuh, and Reinhold Schneider. Optimization problems in contracted tensor networks. *Computing and Visualization in Science*, 14(6):271–285, 2011.

[66] Mike Espig, Wolfgang Hackbusch, Thorsten Rohwedder, and Reinhold Schneider. Variational calculus with sums of elementary tensors of fixed rank. *Numerische Mathematik*, 2012.

[67] G. Evenbly and G. Vidal. Tensor network states and geometry. *Journal of Statistical Physics*, 145(4):891–918, 2011.

[68] Antonio Falcó, Wolfgang Hackbusch, and Anthony Nouy. Geometric structures in tensor representations (release 2). Preprint, Max Planck Institute for Mathematics in the Sciences, 2014.

[69] Andrew J. Ferris. Area law and real-space renormalization. *Phys. Rev. B*, 87:125139, Mar 2013.

[70] E. Fertitta, B. Paulus, G. Barcza, and Legeza Ö. Investigation of metal-insulator like transition through the ab initio density matrix renormalization group approach. *arXiv [cond-mat.str-el]*, page 1406.7038, 2014.

[71] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[72] Miroslav Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematical Journal*, 25(4):619–633, 1975.

[73] Shigeru Furuichi. Tsallis entropies and their theorems, properties and applications. In *Aspects of Optical Sciences and Quantum Information*, pages 1–86. Research Signpost, 2007.

[74] Debashree Ghosh, Johannes Hachmann, Takeshi Yanai, and Garnet Kin-Lic Chan. Orbital optimization in the density matrix renormalization group, with applications to polyenes and -carotene. *The Journal of Chemical Physics*, 128(14):144117, 2008.

[75] Mark S. Gordon and Michael W. Schmidt. Chapter 41 - advances in electronic structure theory: Gamess a decade later. In Clifford E. DykstraGernot FrenkingKwang S. KimGustavo E. Scuseria, editor, *Theory and Applications of Computational Chemistry*, pages 1167 – 1189. Elsevier, Amsterdam, 2005.

[76] Lars Grasedyck. Hierarchical singular value decomposition of tensors. *SIAM Journal on Matrix Analysis and Applications*, 31(4):2029–2054, 2010.

[77] Lars Grasedyck and Wolfgang Hackbusch. An introduction to hierarchical (H-) rank and tt-rank of tensors with examples. *Computational methods in applied mathematics*, 11(3):291–304, 2011.

[78] Lars Grasedyck, Daniel Kressner, and Christine Tobler. A literature survey of low-rank tensor approximation techniques. *GAMM-Mitteilungen*, 36(1):53–78, 2013.

[79] Johannes Hachmann, Wim Cardoen, and Garnet Kin-Lic Chan. Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group. *The Journal of Chemical Physics*, 125(14):144101, 2006.

[80] Johannes Hachmann, Jonathan J. Dorando, Michael Avilés, and Garnet Kin-Lic Chan. The radical character of the acenes: A density matrix renormalization group study. *The Journal of Chemical Physics*, 127(13):134309, 2007.

[81] W. Hackbusch and S. Kühn. A new scheme for the tensor representation. *Journal of Fourier Analysis and Applications*, 15(5):706–722, 2009.

[82] Wolfgang Hackbusch. *Tensor spaces and numerical tensor calculus*, volume 42 of *Springer series in computational mathematics*. Springer, Heidelberg, 2012.

[83] Wolfgang Hackbusch. Numerical tensor calculus. *Acta numerica*, 23:651–742, 2014.

[84] Wolfgang Hackbusch and Reinhold Schneider. Tensor spaces and hierarchical tensor representations. In *Extraction of Quantifiable Information from Complex Systems*, Lecture notes in computational science and engineering. Springer, 2014.

[85] Jutho Haegeman. *Variational renormalization group methods for extended quantum systems*. PhD thesis, Ghent University, 2011.

[86] Jutho Haegeman, J. Ignacio Cirac, Tobias J. Osborne, Iztok Pižorn, Henri Verschelde, and Frank Verstraete. Time-dependent variational principle for quantum lattices. *Phys. Rev. Lett.*, 107:070601, Aug 2011.

[87] Jutho Haegeman, Michaël Mariën, Tobias J Osborne, and Frank Verstraete. Geometry of matrix product states: metric, parallel transport, and curvature. *Journal of Mathematical Physics*, 55(2):50, 2014.

[88] Jutho Haegeman, Tobias J. Osborne, and Frank Verstraete. Post-matrix product state methods: To tangent space and beyond. *Phys. Rev. B*, 88:075133, Aug 2013.

[89] Jutho Haegeman, Bogdan Pirvu, David J. Weir, J. Ignacio Cirac, Tobias J. Osborne, Henri Verschelde, and Frank Verstraete. Variational matrix product ansatz for dispersion relations. *Phys. Rev. B*, 85:100408, Mar 2012.

[90] G. Hager, E. Jeckelmann, H. Fehske, and G. Wellein. Parallelization strategies for density matrix renormalization group algorithms on shared-memory systems. *Journal of Computational Physics*, 194(2):795 – 808, 2004.

[91] Ernst Hairer, Christian Lubich, and Gerhard Wanner. Geometric numerical integration illustrated by the störmer/verlet method. *Acta Numerica*, 12:399–450, 2003.

[92] Karen Hallberg. Density matrix renormalization. In David Sénéchal, André-Marie Tremblay, and Claude Bourbonnais, editors, *Theoretical Methods for Strongly Correlated Electrons*, CRM Series in Mathematical Physics, pages 3–37. Springer New York, 2004.

[93] Karen A. Hallberg. New trends in density matrix renormalization. *Advances in Physics*, 55(5-6):477–526, 2006.

[94] Travis V. Harris, Yuki Kurashige, Takeshi Yanai, and Keiji Morokuma. Ab initio density matrix renormalization group study of magnetic coupling in dinuclear iron and chromium complexes. *The Journal of Chemical Physics*, 140(5):054303, 2014.

[95] M. B. Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, 2007.

[96] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. Wiley New York, 2000.

[97] Christopher J. Hillar and Lek-Heng Lim. Most tensor problems are np-hard. *J. ACM*, 60(6):45:1–45:39, November 2013.

[98] Walter Hofstetter. Generalized numerical renormalization group for dynamical quantities. *Phys. Rev. Lett.*, 85:1508–1511, Aug 2000.

[99] S. Holtz, T. Rohwedder, and R. Schneider. The alternating linear scheme for tensor optimization in the tensor train format. *SIAM Journal on Scientific Computing*, 34(2):A683–A713, 2012.

[100] Sebastian Holtz, Thorsten Rohwedder, and Reinhold Schneider. On manifolds of tensors of fixed tt-rank. *Numerische Mathematik*, 120(4):701–731, 2012.

[101] Michał Horodecki. Entanglement measures. *Quant. Inf. Comp.*, 1:3, May 2001.

[102] R. Horodecki, P. Horodecki, and M. Horodecki. Quantum α-entropy inequalities: independent condition for local realism? *Physics Letters A*, 210(6):377 – 381, 1996.

[103] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009.

[104] Frank Jensen. *Introduction to Computational Chemistry*. Wiley, New York, 1999.

[105] Tamás Juhász and David A. Mazziotti. The cumulant two-particle reduced density matrix as a measure of electron correlation and entanglement. *The Journal of Chemical Physics*, 125(17):174105, 2006.

[106] R. Jullien, P. Pfeuty, J. N. Fields, and S. Doniach. Zero-temperature renormalization method for quantum systems. i. ising model in a transverse field in one dimension. *Phys. Rev. B*, 18:3568–3578, Oct 1978.

[107] Leo P. Kadanoff. Scaling laws for ising models near tc. *Physics*, 2:263, 1966.

[108] Sebastian F. Keller and Markus Reiher. Determining factors for the accuracy of dmrg in chemistry. *CHIMIA International Journal for Chemistry*, 68(4):200–203, 2014-04-30T00:00:00.

[109] Stefan Knecht, Örs Legeza, and Markus Reiher. Communication: Four-component density matrix renormalization group. *The Journal of Chemical Physics*, 140(4):041101, 2014.

[110] Tamara G. Kolda and Brett W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[111] Yuki Kurashige. Multireference electron correlation methods with density matrix renormalisation group reference functions. *Molecular Physics*, 112(11):1485–1494, 2014.

[112] Yuki Kurashige, Garnet Kin-Lic Chan, and Takeshi Yanai. Entangled quantum electronic wavefunctions of the Mn_4CaO_5 cluster in photosystem II. *Nature Chemistry*, 5:660–666, 2013.

[113] Yuki Kurashige, Masaaki Saitow, Jakub Chalupsky, and Takeshi Yanai. Radical o-o coupling reaction in diferrate-mediated water oxidation studied using multireference wave function theory. *Phys. Chem. Chem. Phys.*, 16:11988–11999, 2014.

[114] Yuki Kurashige and Takeshi Yanai. High-performance ab initio density matrix renormalization group method: Applicability to large-scale multireference problems for metal compounds. *The Journal of Chemical Physics*, 130(23):234114, 2009.

[115] Yuki Kurashige and Takeshi Yanai. Second-order perturbation theory with a density matrix renormalization group self-consistent field reference function: Theory and application to the study of chromium dimer. *The Journal of Chemical Physics*, 135(9):094104, 2011.

[116] Nicolas Laflorencie, Erik S. Sørensen, Ming-Shyang Chang, and Ian Affleck. Boundary effects in the critical scaling of entanglement entropy in 1d systems. *Phys. Rev. Lett.*, 96:100603, Mar 2006.

[117] Tran Nguyen Lan, Yuki Kurashige, and Takeshi Yanai. Toward reliable prediction of hyperfine coupling constants using ab initio density matrix renormalization group method: Diatomic $^2\sigma$ and vinyl radicals as test cases. *Journal of Chemical Theory and Computation*, 10(5):1953–1967, 2014.

[118] Cornelius Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *Journal of Research of the National Bureau of Standards*, 45:255, 1950.

[119] J.M. Landsberg. *Tensors: Geometry and Applications*. Graduate studies in mathematics. American Mathematical Society, 2012.

[120] Joseph M. Landsberg, Yang Qi, and Ke Ye. On the geometry of tensor network states. *Quantum Info. Comput.*, 12(3-4):346–354, March 2012.

[121] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, March 2000.

[122] J.M. Lee. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer, 2003.

[123] Ö. Legeza, C. P. Moca, A. I. Toth, I. Weymann, and G. Zarand. Manual for the flexible dm-nrg code. *arXiv [cond-mat.str-el]*, page 0809.3143, 2008.

[124] Ö. Legeza, R.M. Noack, J. Sólyom, and L. Tincani. Applications of quantum information in the density-matrix renormalization group. In H. Fehske, R. Schneider, and A. Weie, editors, *Computational Many-Particle Physics*, volume 739 of *Lecture Notes in Physics*, pages 653–664. Springer Berlin Heidelberg, 2008.

[125] Ö. Legeza, J. Röder, and B. A. Hess. Controlling the accuracy of the density-matrix renormalization-group method: The dynamical block state selection approach. *Phys. Rev. B*, 67:125114, Mar 2003.

[126] Ö. Legeza, J. Röder, and B. A. Hess. Qc-dmrg study of the ionic-neutral curve crossing of LiF. *Molecular Physics*, 101(13):2019–2028, 2003.

[127] Ö. Legeza and J. Sólyom. Optimizing the density-matrix renormalization group method using quantum information entropy. *Phys. Rev. B*, 68:195116, Nov 2003.

[128] Ö. Legeza and J. Sólyom. Optimizing density-matrix renormalization group method using quantum information entropy. In *International Workshop on Recent Progress and Prospects in Density-Matrix Renormalization*. Lorentz Center, Leiden University, The Netherlands, 2004.

[129] Ö. Legeza and J. Sólyom. Quantum data compression, quantum information generation, and the density-matrix renormalization-group method. *Phys. Rev. B*, 70:205118, Nov 2004.

[130] Ö. Legeza and J. Sólyom. Two-site entropy and quantum phase transitions in low-dimensional models. *Phys. Rev. Lett.*, 96:116401, Mar 2006.

[131] Ö. Legeza, J. Sólyom, L. Tincani, and R. M. Noack. Entropic analysis of quantum phase transitions from uniform to spatially inhomogeneous phases. *Phys. Rev. Lett.*, 99:087203, Aug 2007.

[132] Örs Legeza and Gábor Fáth. Accuracy of the density-matrix renormalization-group method. *Phys. Rev. B*, 53:14349–14358, Jun 1996.

[133] Örs Legeza, Florian Gebhard, and Jörg Rissler. Entanglement production by independent quantum channels. *Phys. Rev. B*, 74:195112, Nov 2006.

[134] Örs Legeza, Catalin Pascu Moca, Anna Tóth, Ireneusz Weymann, and ZarándGergely. Flexible dm-nrg, `http://www.phy.bme.hu/~dmnrg/`, 2008.

[135] Örs Legeza, Thorsten Röhwedder, and Reinhold Schneider. Numerical approaches for high-dimensional pdes for quantum chemistry. In Björn Engquist, T. Chan, W.J. Cook, E. Hairer, J. Hastad, A. Iserles, H.P. Langtangen, C. Le Bris, P.L. Lions, C. Lubich, A.J. Majda, J. McLaughlin, R.M. Nieminen, J. ODEN, P. Souganidis, and A. Tveito, editors, *Encyclopedia of Applied and Computational Mathematics*. Springer, 2013.

[136] Örs Legeza, Thorsten Rohwedder, Reinhold Schneider, and Szilárd Szalay. Tensor product approximation (dmrg) and coupled cluster method in quantum chemistry. In Volker Bach and Luigi Delle Site, editors, *Many-Electron Approaches in Physics, Chemistry and Mathematics*, Mathematical Physics Studies, pages 53–76. Springer International Publishing, 2014.

[137] Örs Legeza and Jenö Sólyom. Stability of the haldane phase in anisotropic magnetic ladders. *Phys. Rev. B*, 56:14449–14455, Dec 1997.

[138] Zhening Li, André Uschmajew, and Shuzhong Zhang. On convergence of the maximum block improvement method. *preprint*, sep 2013. Revised April 2014.

[139] Fengyi Liu, Yuki Kurashige, Takeshi Yanai, and Keiji Morokuma. Multireference ab initio density matrix renormalization group (DMRG)-CASSCF and DMRG-CASPT2 study on the photochromic ring opening of spiropyran. *Journal of Chemical Theory and Computation*, 9(10):4462–4469, 2013.

[140] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation by hierarchical tucker and tensor-train tensors. *SIAM Journal on Matrix Analysis and Applications*, 34(2):470–494, 2013.

[141] Christian Lubich. *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*, volume 12 of *Zurich Lectures in Advanced Mathematics*. European Mathematical Society, 2008.

[142] Christian Lubich and I. V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *arXiv [math.NA]*, page 1301.1058, 2013.

[143] Christian Lubich, Ivan Oseledets, and Bart Vandereycken. Time integration of tensor trains. *arXiv preprint arXiv:1407.2042*, 2014.

[144] H.-G. Luo, M.-P. Qin, and T. Xiang. Optimizing Hartree-Fock orbitals by the density-matrix renormalization group. *Phys. Rev. B*, 81:235129, Jun 2010.

[145] A. V. Luzanov and O.V. Prezhdo. High-order entropy measures and spin-free quantum entanglement for molecular problems. *Molecular Physics*, 105(19-22):2879–2891, 2007.

[146] Yingjin Ma and Haibo Ma. Assessment of various natural orbitals as the basis of large active space density-matrix renormalization group calculations. *The Journal of Chemical Physics*, 138(22):224105, 2013.

[147] Konrad H Marti, Bela Bauer, Markus Reiher, Matthias Troyer, and Frank Verstraete. Complete-graph tensor network states: a new fermionic wave function ansatz for molecules. *New Journal of Physics*, 12(10):103008, 2010.

[148] Konrad H. Marti, Irina Malkin Ondík, Gerrit Moritz, and Markus Reiher. Density matrix renormalization group calculations on relative energies of transition metal complexes and clusters. *The Journal of Chemical Physics*, 128(1):014104, 2008.

[149] Konrad H. Marti and Markus Reiher. The density matrix renormalization group algorithm in quantum chemistry. *Zeitschrift für Physikalische Chemie*, 224:583–599, 2010.

[150] Konrad H. Marti and Markus Reiher. Dmrg control using an automated Richardson-type error protocol. *Molecular Physics*, 108(3-4):501–512, 2010.

[151] Konrad H. Marti and Markus Reiher. New electron correlation theories for transition metal chemistry. *Phys. Chem. Chem. Phys.*, 13:6750–6759, 2011.

[152] I. P. McCulloch and M. Gulácsi. The non-abelian density matrix renormalization group algorithm. *EPL (Europhysics Letters)*, 57(6):852, 2002.

[153] Ian P McCulloch. From density-matrix renormalization group to matrix product states. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(10):P10014, 2007.

[154] Ian P. McCulloch and Miklós Gulácsi. Density matrix renormalisation group method and symmetries of the hamiltonian. *Australian Journal of Physics*, 53:597–612, 2000.

[155] Ian P. McCulloch and Miklós Gulácsi. Total spin in the density matrix renormalization group algorithm. *Philosophical Magazine Letters*, 81(6):447–453, 2001.

[156] HD Meyer, Uwe Manthe, and LS Cederbaum. The multi-configurational time-dependent Hartree approach. *CRchemical Physics Letters*, 165(1):73–78, 1990.

[157] Alexander O. Mitrushchenkov, Guido Fano, Roberto Linguerri, and Paolo Palmieri. On the importance of orbital localization in qc-dmrg calculations. *International Journal of Quantum Chemistry*, 112(6):1606–1619, 2012.

[158] A. O. Mitrushenkov, Roberto Linguerri, Paolo Palmieri, and Guido Fano. Quantum chemistry using the density matrix renormalization group ii. *The Journal of Chemical Physics*, 119(8):4148–4158, 2003.

[159] Alexander O. Mitrushenkov, Guido Fano, Fabio Ortolani, Roberto Linguerri, and Paolo Palmieri. Quantum chemistry using the density matrix renormalization group. *The Journal of Chemical Physics*, 115(15):6815–6821, 2001.

[160] Wataru Mizukami, Yuki Kurashige, and Takeshi Yanai. Communication: Novel quantum states of electron spins in polycarbenes from ab initio density matrix renormalization group calculations. *The Journal of Chemical Physics*, 133(9):091101, 2010.

[161] Wataru Mizukami, Yuki Kurashige, and Takeshi Yanai. More electrons make a difference: Emergence of many radicals on graphene nanoribbons studied by ab initio dmrg theory. *Journal of Chemical Theory and Computation*, 9(1):401–407, 2013.

[162] Gerrit Moritz, Bernd Artur Hess, and Markus Reiher. Convergence behavior of the density-matrix renormalization group algorithm for optimized orbital orderings. *The Journal of Chemical Physics*, 122(2):024107, 2005.

[163] Gerrit Moritz and Markus Reiher. Construction of environment states in quantum-chemical density-matrix renormalization group calculations. *The Journal of Chemical Physics*, 124(3):034103, 2006.

[164] Gerrit Moritz and Markus Reiher. Decomposition of density matrix renormalization group states into a slater determinant basis. *The Journal of Chemical Physics*, 126(24):244109, 2007.

[165] Gerrit Moritz, Alexander Wolf, and Markus Reiher. Relativistic dmrg calculations on the curve crossing of cesium hydride. *The Journal of Chemical Physics*, 123(18):184105, 2005.

[166] Matthieu Mottet, Pawel Tecmer, Katharina Boguslawski, Ors Legeza, and Markus Reiher. Quantum entanglement in carbon-carbon, carbon-phosphorus and silicon-silicon bonds. *Phys. Chem. Chem. Phys.*, 16:8872–8880, 2014.

[167] V. Murg, F. Verstraete, and J. I. Cirac. Variational study of hard-core bosons in a two-dimensional optical lattice using projected entangled pair states. *Phys. Rev. A*, 75:033605, Mar 2007.

[168] V. Murg, F. Verstraete, and J. I. Cirac. Exploring frustrated spin systems using projected entangled pair states. *Phys. Rev. B*, 79:195119, May 2009.

[169] V. Murg, F. Verstraete, Ö. Legeza, and R. M. Noack. Simulating strongly correlated quantum systems with tree tensor networks. *Phys. Rev. B*, 82:205105, Nov 2010.

[170] V. Murg, F Verstraete, R Schneider, P. R. Nagy, and Ö Legeza. Tree tensor network state study of the ionic-neutral curve crossing of LiF. *arXiv [physics.chem-ph]*, page 1403.0981, 2014.

[171] Á. Nagy and Robert G. Parr. Information entropy as a measure of the quality of an approximate electronic wave function. *International Journal of Quantum Chemistry*, 58(4):323–327, 1996.

[172] Naoki Nakatani and Garnet Kin-Lic Chan. Efficient tree tensor network states (ttns) for quantum chemistry: Generalizations of the density matrix renormalization group algorithm. *The Journal of Chemical Physics*, 138(13):134113, 2013.

[173] Naoki Nakatani, Sebastian Wouters, Dimitri Van Neck, and Garnet Kin-Lic Chan. Linear response theory for the density matrix renormalization group: Efficient algorithms for strongly correlated excited states. *The Journal of Chemical Physics*, 140(2):024108, 2014.

[174] Roman F. Nalewajski. Entropic measures of bond multiplicity from the information theory. *The Journal of Physical Chemistry A*, 104(51):11940–11951, 2000.

[175] Csaba Nemes, Gergely Barcza, Zoltn Nagy, rs Legeza, and Pter Szolgay. The density matrix renormalization group algorithm on kilo-processor architectures: Implementation and trade-offs. *Computer Physics Communications*, 185(6):1570 – 1581, 2014.

[176] Eric Neuscamman, Takeshi Yanai, and Garnet Kin-Lic Chan. Strongly contracted canonical transformation theory. *The Journal of Chemical Physics*, 132(2):024106, 2010.

[177] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 1 edition, October 2000.

[178] Reinhard M. Noack and Salvatore R. Manmana. Diagonalization and numerical renormalizationgroupbased methods for interacting quantum systems. *AIP Conference Proceedings*, 789(1):93–163, 2005.

[179] Masanori Ohya and Dénes Petz. *Quantum Entropy and Its Use*. Springer Verlag, 1 edition, October 1993.

[180] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349(0):117 – 158, 2014.

[181] I. Oseledets. Approximation of $2^d \times 2^d$ matrices using tensor decomposition. *SIAM Journal on Matrix Analysis and Applications*, 31(4):2130–2145, 2010.

[182] I. Oseledets. Tensor-train decomposition. *SIAM Journal on Scientific Computing*, 33(5):2295–2317, 2011.

[183] I. V. Oseledets and E. E. Tyrtyshnikov. Breaking the curse of dimensionality, or how to use SVD in many dimensions. *SIAM J. Sci. Comput.*, 31(5):3744–3759, 2009.

[184] Stellan Östlund and Stefan Rommer. Thermodynamic limit of density matrix renormalization. *Phys. Rev. Lett.*, 75:3537–3540, Nov 1995.

[185] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. Matrix product state representations. *Quantum Info. Comput.*, 7(5):401–430, jul 2007.

[186] I. Peschel, X. Wang, M. Kaulke, and K. Hallberg, editors. *Density-Matrix Renormalization - A New Numerical Method in Physics*, volume 528 of *Lecture Notes in Physics*. Springer, 1999.

[187] Ingo Peschel, Matthias Kaulke, and rs Legeza. Density-matrix spectra for integrable models. *Annalen der Physik*, 8(2):153–164, 1999.

[188] Robert Peters, Thomas Pruschke, and Frithjof B. Anders. Numerical renormalization group approach to Green's functions for quantum impurity models. *Phys. Rev. B*, 74:245114, Dec 2006.

[189] Dénes Petz. *Quantum Information Theory and Quantum Statistics*. Springer, 2008.

[190] B. Pirvu, J. Haegeman, and F. Verstraete. Matrix product state based algorithm for determining dispersion relations of quantum spin chains with periodic boundary conditions. *Phys. Rev. B*, 85:035130, Jan 2012.

[191] B Pirvu, V Murg, J I Cirac, and F Verstraete. Matrix product operator representations. *New Journal of Physics*, 12(2):025012, 2010.

[192] S. Pittel and N. Sandulescu. Density matrix renormalization group and the nuclear shell model. *Phys. Rev. C*, 73:014301, Jan 2006.

[193] M. B. Plenio, J. Eisert, J. Dreißig, and M. Cramer. Entropy, entanglement, and area: Analytical results for harmonic lattice systems. *Phys. Rev. Lett.*, 94:060503, Feb 2005.

[194] Maren Podewitz, Martin T. Stiebritz, and Markus Reiher. An enquiry into theoretical bioinorganic chemistry: How heuristic is the character of present-day quantum chemical methods? *Faraday Discuss.*, 148:119–135, 2011.

[195] John Preskill. Lecture notes on quantum computation. available on the homepage of the author: `http://www.theory.caltech.edu/people/preskill/ph229`.

[196] Alan Race, William Barford, and Robert J. Bursill. Density matrix renormalization calculations of the relaxed energies and solitonic structures of polydiacetylene. *Phys. Rev. B*, 67:245202, Jun 2003.

[197] Michael Reed and Barry Simon. *Methods of modern mathematical physics. IV. , Analysis of operators*. Academic Press, London, 1978.

[198] Markus Reiher. On the definition of local spin in relativistic and nonrelativistic quantum chemistry. *Faraday Discuss.*, 135:97–124, 2007.

[199] Markus Reiher. A theoretical challenge: Transition-metal compounds. *CHIMIA International Journal for Chemistry*, 63(3):140–145, 2009-03-25T00:00:00.

[200] Julián Rincón, D.J. García, and K. Hallberg. Improved parallelization techniques for the density matrix renormalization group. *Computer Physics Communications*, 181(8):1346 – 1351, 2010.

[201] Jörg Rissler, Reinhard M. Noack, and Steven R. White. Measuring orbital interaction using quantum information theory. *Chemical Physics*, 323(23):519 – 531, 2006.

[202] Michael Rohlfing and Steven G. Louie. Optical excitations in conjugated polymers. *Phys. Rev. Lett.*, 82:1959–1962, Mar 1999.

[203] T. Rohwedder and A. Uschmajew. On local convergence of alternating schemes for optimization of convex problems in the tensor train format. *SIAM Journal on Numerical Analysis*, 51(2):1134–1162, 2013.

[204] Thorsten Rohwedder and Reinhold Schneider. Error estimates for the coupled cluster method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47:1553–1582, 11 2013.

[205] J. Rotureau, N. Michel, W. Nazarewicz, M. Płoszajczak, and J. Dukelsky. Density matrix renormalization group approach for many-body open quantum systems. *Phys. Rev. Lett.*, 97:110603, Sep 2006.

[206] Masaaki Saitow, Yuki Kurashige, and Takeshi Yanai. Multireference configuration interaction theory using cumulant reconstruction with internal contraction of density matrix renormalization group wave function. *The Journal of Chemical Physics*, 139(4):044118, 2013.

[207] Niyasi Serdar Sariciftci, editor. *Primary Photoexcitations in Conjugated Polymers: Molecular Exciton Versus Semiconductor Band Model*. World Scientific, 1998.

[208] D. V. Savostyanov, S. V. Dolgov, J. M. Werner, and Ilya Kuprov. Exact NMR simulation of protein-size spin systems using tensor train formalism. *Phys. Rev. B*, 90:085139, Aug 2014.

[209] Erhard Schmidt. Zur theorie der linearen und nichtlinearen integralgleichungen. i. teil: Entwicklung willkürlicher funktionen nach systemen vorgeschriebener. *Mathematische Annalen*, 63:433–476, 1907.

[210] Reinhold Schneider and Andr Uschmajew. Approximation rates for the hierarchical tensor format in periodic sobolev spaces. *J. Complexity*, pages 56–71, 2014.

[211] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77:259–315, Apr 2005.

[212] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96 – 192, 2011. January 2011 Special Issue.

[213] Michel Schott. Optical properties of single conjugated polymer chains (polydiacetylenes). In Guglielmo Lanzani, editor, *Photophysics of Molecular Materials: From Single Molecules to Single Crystals*, pages 49–151. Wiley-VCH Verlag GmbH & Co. KGaA, 2006.

[214] Ervin Schrödinger. Die gegenwrtige situation in der quantenmechanik. *Naturwissenschaften*, 23:807, 1935.

[215] Ervin Schrödinger. Discussion of probability relations between separated systems. *Math. Proc. Camb. Phil. Soc.*, 31:555, 1935.

[216] Sandeep Sharma and Garnet Kin-Lic Chan. Block code for dmrg, `http://www.princeton.edu/chemistry/chan/software/dmrg/`, 2012.

[217] Sandeep Sharma and Garnet Kin-Lic Chan. Spin-adapted density matrix renormalization group algorithms for quantum chemistry. *The Journal of Chemical Physics*, 136(12):124121, 2012.

[218] Sandeep Sharma, Takeshi Yanai, George H. Booth, C. J. Umrigar, and Garnet Kin-Lic Chan. Spectroscopic accuracy directly from quantum chemistry: Application to ground and excited states of beryllium dimer. *The Journal of Chemical Physics*, 140(10):104112, 2014.

[219] Y.-Y. Shi, L.-M. Duan, and G. Vidal. Classical simulation of quantum many-body systems with a tree tensor network. *Phys. Rev. A*, 74:022320, Aug 2006.

[220] Germán Sierra and Tomotoshi Nishino. The density matrix renormalization group method applied to interaction round a face hamiltonians. *Nuclear Physics B*, 495(3):505 – 532, 1997.

[221] Sukhwinder Singh, Robert N. C. Pfeifer, and Guifré Vidal. Tensor network decompositions in the presence of a global symmetry. *Phys. Rev. A*, 82:050301, Nov 2010.

[222] Sukhwinder Singh and Guifre Vidal. Tensor network states and algorithms in the presence of a global su(2) symmetry. *Phys. Rev. B*, 86:195114, Nov 2012.

[223] Sukhwinder Singh, Huan-Qiang Zhou, and Guifre Vidal. Simulation of one-dimensional quantum systems with a global su(2) symmetry. *New Journal of Physics*, 12:033029, 2010.

[224] S. Spagnoli, J. Berrhar, C. LapersonneMeyer, and M. Schott. Polydiacetylene chains diluted in their singlecrystal monomer matrix. *The Journal of Chemical Physics*, 100(9):6195–6202, 1994.

[225] E. M. Stoudenmire and Steven R. White. Real-space parallel density matrix renormalization group. *Phys. Rev. B*, 87:155137, Apr 2013.

[226] A Szabo and N S Ostlund. *Modern Quantum Chemistry*. McGraw Hill, New York, 1982.

[227] Szilárd Szalay. Quantum entanglement in finite-dimensional hilbert spaces. *arXiv [quant-ph]*, page 1302.4654, 2013.

[228] L. Tagliacozzo, G. Evenbly, and G. Vidal. Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law. *Phys. Rev. B*, 80:235127, Dec 2009.

[229] Pawel Tecmer, Katharina Boguslawski, Ors Legeza, and Markus Reiher. Unravelling the quantum-entanglement effect of noble gas coordination on the spin ground state of cuo. *Phys. Chem. Chem. Phys.*, 16:719–727, 2014.

[230] B. Thakur, S. Pittel, and N. Sandulescu. Density matrix renormalization group study of ^{48}Cr and ^{56}Ni. *Phys. Rev. C*, 78:041303, Oct 2008.

[231] A. I. Tóth, C. P. Moca, Ö. Legeza, and G. Zaránd. Density matrix numerical renormalization group for non-abelian symmetries. *Phys. Rev. B*, 78:245109, Dec 2008.

[232] André Uschmajew and Bart Vandereycken. The geometry of algorithms using hierarchical tensors. *Linear Algebra and its Applications*, 439(1):133 – 166, 2013.

[233] André Uschmajew and Bart Vandereycken. Line-search methods and rank increase on low-rank matrix varieties. *preprint*, juli 2014. Accepted for NOLTA2014.

[234] Karel Van Acoleyen, Michaël Mariën, and Frank Verstraete. Entanglement rates and area laws. *Phys. Rev. Lett.*, 111:170501, Oct 2013.

[235] J.-W. van der Horst, P. A. Bobbert, M. A. J. Michels, G. Brocks, and P. J. Kelly. Ab initio. *Phys. Rev. Lett.*, 83:4413–4416, Nov 1999.

[236] F. Verstraete and J. I. Cirac. Renormalization algorithms for quantum-many body systems in two and higher dimensions. *arXiv [cond-mat.str-el]*, page 0407066, 2004.

[237] F. Verstraete, V. Murg, and J.I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, 2008.

[238] F. Verstraete, D. Porras, and J. I. Cirac. Density matrix renormalization group and periodic boundary conditions: A quantum information perspective. *Phys. Rev. Lett.*, 93:227205, Nov 2004.

[239] G. Vidal. Entanglement renormalization. *Phys. Rev. Lett.*, 99:220405, Nov 2007.

[240] G. Vidal. Class of quantum many-body states that can be efficiently simulated. *Phys. Rev. Lett.*, 101:110501, Sep 2008.

[241] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev. Entanglement in quantum critical phenomena. *Phys. Rev. Lett.*, 90:227902, Jun 2003.

[242] Guifré Vidal. Entanglement monotones. *Journal of Modern Optics*, 47(2-3):355–376, 2000.

[243] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91:147902, Oct 2003.

[244] Haobin Wang and Michael Thoss. Multilayer formulation of the multiconfiguration time-dependent Hartree theory. *The Journal of Chemical Physics*, 119(3):1289–1299, 2003.

[245] Andreas Weichselbaum. Non-abelian symmetries in tensor networks: A quantum symmetry space approach. *Annals of Physics*, 327(12):2972 – 3047, 2012.

[246] Andreas Weichselbaum and Jan von Delft. Sum-rule conserving spectral functions from the numerical renormalization group. *Phys. Rev. Lett.*, 99:076402, Aug 2007.

[247] H. Weyl. Quantenmechanik und gruppentheorie. *Zeitschrift für Physik*, 46(1-2):1–46, 1927.

[248] S. R. White and R. M. Noack. Real-space quantum renormalization groups. *Phys. Rev. Lett.*, 68:3487–3490, Jun 1992.

[249] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[250] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, Oct 1993.

[251] Steven R. White. Numerical canonical transformation approach to quantum many-body problems. *The Journal of Chemical Physics*, 117(16):7472–7482, 2002.

[252] Steven R. White. Density matrix renormalization group algorithms with a single center site. *Phys. Rev. B*, 72:180403, Nov 2005.

[253] Steven R. White and Richard L. Martin. Ab initio quantum chemistry using the density matrix renormalization group. *The Journal of Chemical Physics*, 110(9):4127–4130, 1999.

[254] Eugene Paul Wigner. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*. Springer, 1931.

[255] Eugene Paul Wigner. On unitary representations of the inhomogeneous lorentz group. *Annals of Mathematics*, 40(1):pp. 149–204, 1939.

[256] Eugene Paul Wigner. *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*. Academic Press, 1959.

[257] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013. Cambridge Books Online.

[258] Kenneth G. Wilson. The renormalization group: Critical phenomena and the kondo problem. *Rev. Mod. Phys.*, 47:773–840, Oct 1975.

[259] Sebastian Wouters. Accurate variational electronic structure calculations with the density matrix renormalization group. *arXiv [cond-mat.str-el]*, page 1405.1225, 2014. PhD thesis, Ghent University.

[260] Sebastian Wouters. CheMPS2: a spin-adapted implementation of DMRG for ab initio quantum chemistry, `https://github.com/SebWouters/CheMPS2`, 2014.

[261] Sebastian Wouters, Thomas Bogaerts, Pascal Van Der Voort, Veronique Van Speybroeck, and Dimitri Van Neck. Communication: Dmrg-scf study of the singlet, triplet, and quintet states of oxo-mn(salen). *The Journal of Chemical Physics*, 140(24):241103, 2014.

[262] Sebastian Wouters, Peter A. Limacher, Dimitri Van Neck, and Paul W. Ayers. Longitudinal static optical properties of hydrogen chains: Finite field extrapolations of matrix product state calculations. *The Journal of Chemical Physics*, 136(13):134110, 2012.

[263] Sebastian Wouters, Naoki Nakatani, Dimitri Van Neck, and Garnet Kin-Lic Chan. Thouless theorem for matrix product states and subsequent post density matrix renormalization group methods. *Phys. Rev. B*, 88:075122, Aug 2013.

[264] Sebastian Wouters, Ward Poelmans, Paul W. Ayers, and Dimitri Van Neck. CheMPS2: A free open-source spin-adapted implementation of the density matrix renormalization group for ab initio quantum chemistry. *Computer Physics Communications*, 185(6):1501 – 1514, 2014.

[265] Sebastian Wouters and Dimitri Van Neck. The density matrix renormalization group for ab initio quantum chemistry. *The European Physical Journal D*, 68(9), 2014.

[266] T. Xiang. Density-matrix renormalization-group method in momentum space. *Phys. Rev. B*, 53:R10445–R10448, Apr 1996.

[267] Susumu Yamada, Toshiyuki Imamura, and Masahiko Machida. Parallelization design on multi-core platforms in density matrix renormalization group toward 2-d quantum strongly-correlated systems. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '11, pages 62:1–62:10, New York, NY, USA, 2011. ACM.

[268] Takeshi Yanai, Yuki Kurashige, Debashree Ghosh, and Garnet Kin-Lic Chan. Accelerating convergence in iterative solution for large-scale complete active space self-consistent-field calculations. *International Journal of Quantum Chemistry*, 109(10):2178–2190, 2009.

[269] Takeshi Yanai, Yuki Kurashige, Eric Neuscamman, and Garnet Kin-Lic Chan. Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *The Journal of Chemical Physics*, 132(2):024105, 2010.

[270] Takeshi Yanai, Yuki Kurashige, Eric Neuscamman, and Garnet Kin-Lic Chan. Extended implementation of canonical transformation theory: parallelization and a new level-shifted condition. *Phys. Chem. Chem. Phys.*, 14:7809–7820, 2012.

[271] Harry Yserentant. On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numer. Math.*, 98(4):731–759, 2004.

[272] Harry Yserentant. *Regularity and approximability of electronic wave functions*. Lecture Notes in Mathematics. Springer, Berlin, 2010.

[273] J.F. Yu, H.-C. Hsiao, and Ying-Jer Kao. GPU accelerated tensor contractions in the plaquette renormalization scheme. *Computers & Fluids*, 45(1):55 – 58, 2011. 22nd International Conference on Parallel Computational Fluid Dynamics (ParCFD 2010) ParCFD.

[274] Dominika Zgid and Marcel Nooijen. The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space. *The Journal of Chemical Physics*, 128(14):144116, 2008.

[275] Dominika Zgid and Marcel Nooijen. Obtaining the two-body density matrix in the density matrix renormalization group method. *The Journal of Chemical Physics*, 128(14):144115, 2008.

[276] Dominika Zgid and Marcel Nooijen. On the spin and symmetry adaptation of the density matrix renormalization group method. *The Journal of Chemical Physics*, 128(1):014107, 2008.

[277] Paul Ziesche. Correlation strength and information entropy. *International Journal of Quantum Chemistry*, 56(4):363–369, 1995.